

Possible Proto-Asian Archaic Residue and the Statigraphy of Diffusional Cumulation in Austro-Asian Languages¹

Michael D. Larish, Ph.D.
 University of Hawai‘i
 Hawai‘i Community College (Hilo)
 larish@hawaii.edu

Abstract

This speculative paper considers the Proto-Asian hypothesis (**PAsn-Hyp**), viz., most modern languages found in Asia, Southeast Asia, and Austronesia may descend from one mother language called Proto-Asian (**PAsn**).² Thus, Austroasiatic (AA), Austronesian (An), Japanese-Korean, Kam-Sui (KS), Miao-Yao (MY), Sino-Tibetan (ST), and Tai-Kadai (TK) may possibly subgroup together.³ A second purpose of this paper is to begin a

¹ I would like to thank the numerous scholars at 10-ICAL who gave me feedback on an earlier draft of this paper.

² I propose the term Austro-Asian be extended to include the area between Rapanui (Easter Island) and Tibet (East-West) and China and New Zealand (North-South).

³ Abbreviations used and data sources in this paper not found in Larish (1999) include the following:

AA Austroasiatic,
 Asl Aslian [MK],
 An Austronesian,
 JAM James A. Matisoff (2003),
 KS Kam-Sui,
 MY Miao-Yao,
 OC Old Chinese [(Coblin 1986, Matisoff 2003) = Karlgren’s “Archaic Chinese”],
 OJpn Old Japanese (Martin 1987, Benedict 1990),
 PAJ Proto-Austro-Japanese (Benedict 1990),
 PAJ Proto-Austro-Tai (Benedict 1990, 1975),
 PAK Proto-Austro-Kadai (Benedict 1990),
 PAn Proto-Austronesian (Blust 1997),
 PAsn Proto-Asian,
 PAsn-hyp Proto-Asian hypothesis,
 PKB Paul K. Benedict (1990, 1975, 1972),
 PLB Proto-Lolo-Burmese (Matisoff 2003),
 PMACM Proto-Moken-Moklen-Acehnese-Chamic (Larish 1999),
 PMM Proto-Moken-Moklen (Larish 1999),
 PST Proto-Sino-Tibetan (Coblin 1986),
 PTai Proto-Tai (Li 1977),
 PTB Proto-Tibeto-Burman (Matisoff 2003),
 STC Sino-Tibetan: A Conspectus (Benedict 1972),
 TK Tai-Kadai,

methodological and theoretical discussion on the problems of sorting out genetic archaic retention from diffusional cumulation in Austro-Asian languages. The principal data (Appendix 1.1) in this paper were obtained by comparing Proto-Austronesian (**PAn**), Proto-Sino-Tibetan (**PST**), and Proto-Tibeto-Burman (**PTB**) core vocabulary, as reconstructed by Blust (1997), Coblin (1986), and Matisoff (2003), respectively. Sixty-two (62) possible cognates were found in Blust’s revised Swadesh 200-word list. Appendix 2 presents supplemental evidence comparing a number of Austro-Asian languages, incorporating data from Benedict (1990, 1975): Proto-Austro-Thai (**PAT**) and Proto-Austro-Japanese (**PAJ**). The evidence in the appendices tentatively supports the **PASN-Hyp**. It may take decades for the main question raised in this paper (and others) to be answered. Moreover, it will probably be answered—*beyond doubt*—computationally by future comparativists.

1. Introduction

1.1. Overview

This paper is divided into two major sections. Section 1 outlines the theoretical perspective, methods, and references that relate to this ongoing research. Section 2 presents the results of this research, exploring whether the **PASN-Hyp** is workable. Many of subsections that follow this discussion are derived from an unpublished chapter extracted from Larish (1999). They provide specific phonological and lexical examples relating to the lexical affinities between mainland-Austronesian and Mon-Khmer languages; they illustrate some of the problems of distinguishing between archaic genetic retention and more recent diffusional cumulation.⁴

1.2. Theoretical Perspective

Unlike Lexicase Theory, languages are not monostratal in terms of their accumulation of vocabulary over time (see Starosta 1988:2). Moreover, unlike archaeologists, who dig until they reach culturally sterile soil, historical comparativists—historical linguists, cultural anthropologists, historians, and linguistic geographers, for example—have no theoretical time-depth limits unless they are imposed by tradition, basic underlying assumptions, or major theoretical constraints. Linguistic texts, for instance, often include a claim that the historical-comparative method cannot reach past 6,000 to 8,000 years before present, the depth at which recurrent sound correspondences become problematic (cf. O’Grady, Archibald, Aronoff, and Rees-Miller 2005:327). However, proto-language to proto-language comparison—referred to as *long-*

For additional abbreviations, refer to Larish (1999:xxxix).

⁴ These terms were first used by Swadesh (1964) in Hymes’s classic reader.

range comparison or *palaeolinguistics*, extends this limit deeper into the remote past (*cf.* Starostin 1992:76, Trask 2000:66-7).

In the course of researching the **PASN-Hyp**, two new acronyms were developed in order to clarify my thinking about the distinction between direct genetic heritage and the subsequent diffusion between languages that may or may not be genetically related to each other. To put archaic residue and diffusional cumulation into layman's terms, I developed the distinction between Vertical Genetic Retention (\downarrow VGR) and Lateral Loan Relationship (\leftrightarrow LLR). VGR must be determined before LLR can be ascertained. To complicate matters, the historical linguist will be confronted with numerous LLR, from old to recent (see Figure 1 below). The number of LLR depends on how many times language contact has occurred. Language contact can occur between languages with VGR or without VGR. In Asia and Europe, the former case appears to be more common. Consider, for example, the various diffusional influences in English by both Germanic and non-Germanic relatives within Indo-European. Describing the various strata of VGR and LLR is the task of the historical comparativist.

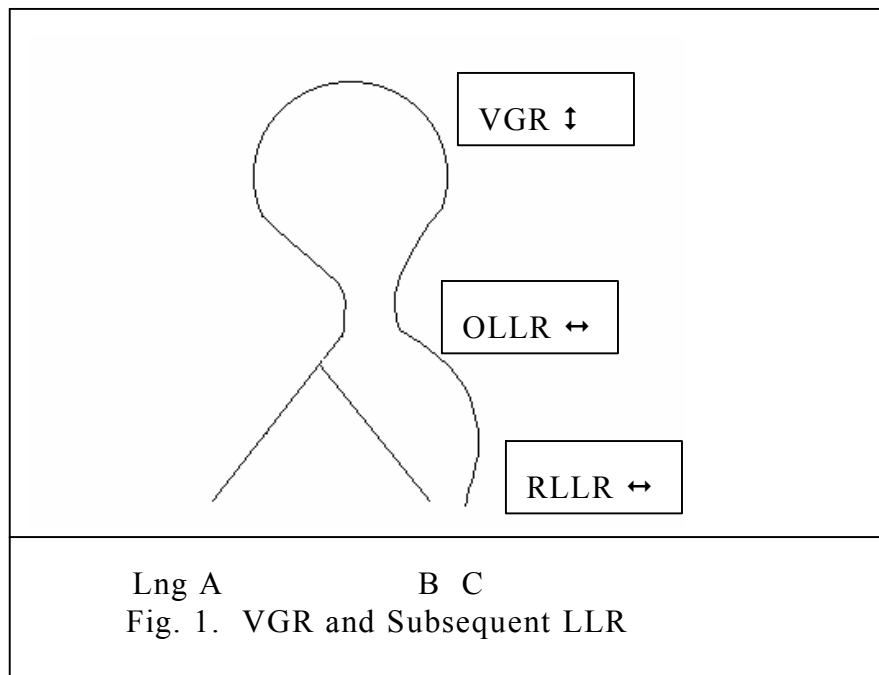


Figure 1 represents two groups of people who separated in a remote period, then drew close again due to migration or population expansion. The neck of the octopus represents an Old Lateral Loan Relationship (\leftrightarrow OLLR) where remote language contact occurs. After Lng A splits from Lng B, speakers of Lng B interact with speakers of Lng C, creating a Recent Lateral Loan Relationship (\leftrightarrow RLLR).

1.3. Methods

The comparanda in Appendix 1 focus on core vocabulary to get to the genetic heart of the **PASn-Hyp**. Ideally, reconstructions with the greatest depth ought to be used in this analysis.⁵ Therefore, **PAn** etyma from Blust (1997) were compared with Proto-Sino-Tibetan (**PST**) forms from Coblin (1986) and Proto-Tibeto-Burman (**PTB**) forms from Matisoff (2003).

Although not explicitly stated, I have followed many of the procedures outlined in Reid's (2005) "The Current Status of Austric." For example, Appendix 1 is separated into section 1.1 (PAn and PST/PTB comparanda) and section 1.2 (PMP and PTB comparanda) since the former will have higher "probative value" than the later. Furthermore, to augment the Swadesh 200 data in Appendix 1, Appendix 2 was included. Appendix 2 mainly compares Benedict's (1975) **PAT** with **PTB** forms (Matisoff 2003); however, some **PAJ** and **PAK** (Benedict 1990) forms are included in Appendix 2 to demonstrate that **OJpn** etyma may also fit into the **PASN** hypothesis. Knowing that Benedict's (1990) **PAJ** and **PAK** reconstructions and his (1975) **PAT** reconstructions are controversial, I have used them for two reasons: (1) Even if future research does not support his reconstructions, Benedict has collected an amazing corpus of hard-to-find languages. (2) **PAT** reconstructions have the time depth needed to support the **PASN-Hyp**. In fact, in many cases, Benedict's (1975) **PAT** reconstructions bring us closer to the **PASN** level. However, the long-held debate about the place of Sino-Tibetan languages in Asian Linguistics must be sustained until a consensus can be reached. Benedict's (1975) point of view is that **PST** does not subgroup with Austro-Thai. In his "Summary and Conclusions for Part 3," (Benedict 1975:116-33, republished from 1967), Benedict's main thesis is that the affinities between ST and AT are the result of "extensive cultural contact between early Chinese and AT peoples" (p. 123) and "that the AT cultural influence extended also far to the west, providing many loan-words (and presumably cultural items) to the TB peoples" (p. 125). Although diffusional cumulation is a likely explanation for some affinities, theoretically speaking, an examination of Vertical Genetic Retention (\downarrow VGR) must be considered prior to hypotheses about Lateral Loan Relationships (\leftrightarrow LLR). As far as I can determine, not having examined all of Benedict's work in his highly productive career, I could find no in-depth comparison/discussion on core vocabulary between PST and PAT. Again, not having examined Sagart's (2005) data and conclusions in Sagart, Blench, and Sanchez-Mazas, I know from Reid's (2005, p. 9) conclusion that Sagart will

⁵ Whenever possible, proto-languages are compared, taking us further back in time. The farther we can push back the depth of the comparison, the better the comparison, for reconstructed etyma are based on a number of languages (at least two). Until reconstructions are available, we remain less sure of genetic or diffusional affinities, for we may be simply comparing lexemes that accidentally look alike

suggest—and has been suggesting for many years—that “With the accumulation of evidence presented by Sagart in this volume and elsewhere, that Austronesian can also be shown to be genetically related to the Sino-Tibetan family of languages.” Therefore, the present paper can be viewed as an independent study that uses different data which may ultimately support Sagart’s conclusions.

1.4. Literature review

In three more years, we will reach the one-hundredth anniversary of Schmidt’s (1909) Austric hypothesis. For review/expansion on the Schmidt’s (1909) Austric hypothesis, see Benedict (1976:1-36), Ruhlen (1991:151-7), and Reid (2005). Reid (1997:19) concludes that ‘we need no longer cautiously refer to this family with the sobriquet “Austric Hypothesis”.’ Professor Reid has also reviewed the work of La Vaughn Hayes in a positive light. Hayes (1992, 1996, 1997, 1999) has been steadily working toward establishing the Austric language family by examining phonological evidence and basic vocabulary. Larish (1999, Appendix C) and Thurgood (1999) have identified a large number of possible Mon-Khmer loans in mainland-An languages.

1.4.1. Benedict (1990, 1976, 1975, 1972)

Benedict (1976) argues for the demise of Austric by suggesting a remote contact relationship between AT and AA. I do not find the following argument convincing: ‘a mainland branch of AT, now extinct, became “substratumized” by AA, yielding up certain roots in the process’ (1976:28). Benedict’s most convincing argument is the general lack of cognates between An and AA in core vocabulary, but Diffloth (1994) counters by proposing that much core vocabulary can be replaced over extreme time depths. Furthermore, Diffloth (1994) suggests that a “probable” genetic AN-MK connection can be observed in the words for *dog*, *fish*, *centipede*, *wood*, *eye*, *bone*, *hair*, *tongue*, and *left*.

Examining PKB’s (1975) PAT reconstructions is difficult yet productive; for example, Benedict (1975:346-8) proposes seven PAT variant reconstructions for ‘open/gape/stand open/force open/force apart/oppose/separate/bay/river’ followed by about two pages of data. Unfortunately, the PAT reconstructions are aligned laterally. When you line them up vertically, it is easier to see the interrelationships to his variant reconstructions (see Appendix 2).

1.4.2. Matisoff (2003)

Matisoff’s (2003) 800-page *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction* is an exemplary corpus of solid scholarship and a primary data source for the present research. It can be obtained on-line: http://repositories.cdlib.org/ucpress/ucpl/vol_135.

2. Remote Austro-Asian Connections

2.1. Introduction

Two possible remote connections between Austro-Asian languages must be considered:

- (i) A vertical genetic relationship between **Pre-An** and other Asian languages
- (ii) Remote contact between **Pre-An** and other Asian “language families” or between possible “branches” of one family

This problem is perhaps best illustrated by the far-reaching lexical agreement for “river” across a number of language families. Matisoff describes the word for “river” as an “East/SE Asian Wanderwort.” Widely distributed lexemes can result from two sources: Vertical Genetic Retention (↓VGR) and Lateral Loan Relationship (↔LLR). Consider the following data:

PAT-75	*[(m)b]aŋaq	‘open/gape/stand open/force open/force	
	*[b/]n/aŋ[aq]	apart/oppose/separate/bay/river’	
	*[(m)ba]ŋa[q/]ŋa[q]	(Benedict 1975:346-8)	
	*[(m)b]akaq		
	*[ba]ga[q/]gaq		
	*[ba]gaŋ(/gaŋ)		
	*[ba]kaŋ(/kaŋ)		
PLB-JAM	*ʔ-ŋak ^L	‘open wide’	(Matisoff 2003:606L)
PLB-JAM	*kaŋ ²	‘spread/stretch out’	(Matisoff 2003:266, 595L)
PTB-JAM	*s-bu	‘open/bud’	(Matisoff 2003:184, 585C, 660R, cf. STC#260)
PTB-JAM	*m-ka	‘open/opening/mouth/door’	(Matisoff 2003:594C, 660R, cf. STC#468)
OC	*g’o	‘door/opening’	(Matisoff 2003:173)
PTB-JAM	*klu(:)ŋ~*k(l)uk	‘valley/river’	(Matisoff 2003:287, 524, 596L, 665L)
PTB-JAM	*klyoŋ~*k(l)uk	‘valley/river’	(Matisoff 2003:294, 596L, 665L)
PLB-JAM	*laŋ ¹	‘stream/river/valley’	(Matisoff 2003:266)
OC	*kŭŋ	‘river’	(Matisoff 2003:287)
PTB-JAM	*klu(:)ŋ~*k(l)uk	‘valley/river’	(Matisoff 2003:287, 524, 596L, 665L)
PTB-JAM	*klyoŋ~*k(l)uk	‘valley/river’	(Matisoff 2003:294, 596L, 665L)
PLB-JAM	*laŋ ¹	‘stream/river/valley’	(Matisoff 2003:266)

- OC** *kǔŋ (Matisoff 2003:287)
 Thai-C k^hlɔːŋ ‘canal’
 [Mon-Khmer] **Proto-Monic** *krooŋ, **PNyK** *krooŋ ‘large river’, **PMon** *krɔ̄iŋ
 ‘stream, creek, small river’
- PMACM** *ŋaŋa(h) ‘to open (mouth)’
PMM *ŋaŋah ‘to open (mouth)’
PPh *ŋǎŋáh ‘open (mouth)’
 Mal ŋaŋa ‘agape’
PMACM *muka[:]?/*buka[:]? ‘to open’
PChm *pǒk revised as *po[:]?/k/h ‘to open’ (< **Pre-Chm**
 ?*po:q) > Rad, Jor pǒk, Rog po? [Chm-Mou pə:h ‘ouvrier (to open)’,
 Chm-Mou bɔː? ‘ouvert (open)’, cf. Chm-Mou prù:h ‘lever (to lift)’]
- PBtk** *uŋkab
PTB-JAM *s-bu ‘open/bud’
 (Matisoff 2003:184, 585C, 660R, cf. STC#260)
PTB-JAM *m-ka ‘open/opening/mouth/door’
 (Matisoff 2003:594C, 660R, cf. STC#468)
PMP *buka (poss. *buká?) ‘to open, uncover’
PMal *buka? > Mal, UL buka
PPh *bǔká? ‘open, v.; opened’
UAN *buka ‘open, v.’
 [Mon-Khmer] **Proto-Monic** *pɔk, **PNyK** *pɔk ‘to open (a door, the lid of a
 pot . . .), to expose (a wound, breasts), to lift (a curtain, a skirt, the
 corner of a mat . . .) (in order to look underneath)’, **PMon** *pɔk ‘to
 open, uncover, expose’
 BM habap-patan, TM pɔk, NyK chlɔk/yùk (TLW) ‘lift’
 Asl-SkBl Tembi, Serau. bu:ka’ (bu:kak) ‘to open’
PAC-Trg ^x*?aha/^x*ha ‘open (mouth to say sthg.)’ (Thurgood 1999:309, §1.2:
MK origin)
PChm *?aha/*ha revised as [?]x(?a)ha ‘to open mouth’ (poss. MK loan) >
 Rad, Jor, Chm ha, Rog ?aha/ha;
 [Mon-Khmer] **Proto-Monic** *(ŋ)haa?, **PNyK** *(ŋ)h_{aa}? “to open one’s mouth,
 or beak” **PMon** *h_a~h_a? ‘to gape open, open wide’
 Nic tom-âng-hala; ong-âng-ha-chakâ “to open one’s mouth”
PJH *ha ‘open mouth’
PSBnr *ha ‘open mouth’
PEK *kah_? ‘open mouth’
 Mkl-BDC k^hlɔːŋ/klɔːŋ (prob. MK loan, poss. indirectly via Thai)
PAC-Trg ^x*krɔːŋ ‘river’ (Thurgood 1999:324, §1.2: **MK Origin**)
 Ach-Saw kruəŋ
PChm *kroːŋ > Rad kroŋ, Jor, Chm krɔŋ [Chm-Mou kroːŋ ‘fleuve’]

Mal	sungai	
Png	ílog	
Tag	ilog	
Thai-C	mê: ná:m	‘river’
Thai-C	k ^h lɔ:ŋ	‘canal’ (prob. MK loan)
[Mon-Khmer]	Proto-Monic *krooŋ, PNyK *kr <u>o</u> oŋ	‘large river’, PMon *kr <u>ʌ</u> iŋ
		‘stream, creek, small river’
BM	krəŋ	‘river (small)’
BM	bi, TM bi, NyK mɛnam (TLW)	‘river (large)’
BM	kləŋ, TM klòŋ, NyK trəw	‘road’ (< ‘waterway?’)
Khm	tùənlè	‘river’
PEK	*kru:ŋ	‘river’

Do you see Vertical Genetic Retention (↓VGR) or Lateral Loan Relationships (↔LLR) in the above data? Matisoff (2003:3) defines “allofams” as “variant forms of the same word-family.” I applaud Benedict’s (1990, 1975) and Matisoff’s (2003) efforts to identify allofams in their data. Rather than provide one or two-word glosses, which is traditional in much historical-comparative research, future computational analysis of such allofams across thousands of Austro-Asian languages will probably play a central role in establishing or rejecting the **PAsn-Hyp**. In speaking with various scholars in Puerto Princesa, Palawan in January 2006, I noted that Benedict’s multi-glossed reconstructions helped me discover connections in my own data that I had not previously noticed. Robert Blust (pers. comm.) asked me for an example of such a discovery. One instance is presented in the data block above. I had not noticed the phonosemantic parallels between the ultimate syllables in the following etyma reconstructed in Larish (1999, Book 2, Appendix C): **PMACM** *ŋaŋa(h) ‘to open (mouth)’ and **PMACM** *muka[:]?/*buka[:]? ‘to open’. This phonosemantic parallel suggests that these etyma may have developed from a common **PAsn** monosyllabic root.

2.2. Results of the present research

Sixty-two (62) possible PAn/PST, PAn/PTB or PMP/PTB cognate sets are presented in Appendix 1.1 and 1.2. Some of these sets will probably be rejected by future research, for I have determined that the Swadesh-200 list is insufficient for establishing recurrent correspondences in proto-language to proto-language comparison. While the Swadesh-200 is sufficient to establish recurrent sound correspondences for closely related languages, the number of comparanda for long-range comparative work must be extended to perhaps 500 core-type sets. Until this work is complete, I will refrain from listing possible correspondences. However, I would like to discuss three possible consonant correspondences. First, **PAn** *R appears to correspond to **PTB** *y in Appendix 1.1 (numbers 1, 9, 16, 20, 25), but other data demonstrate *R to

r* (17, 20, 35) or **R* to **l* (26, 27, 30).⁶ Second, instances of the possible correspondence between **PAn **t* and **PST/PTB** **k* (12, 24, 42, 45) can be noted. Finally, another salient correspondence that appears recurrently in the present research is *k* to *h*. It is found in sets 24 pound/strike, 38 burn, and 42 red in Appendix 1.1. Due to the significance of **PMM** **k* to **PACM** **h* (< **PAn** **q*) in mainland-Austronesian subgrouping arguments (see Larish 1999:363-7, see especially Figure 7.1 and Table 7.1; Larish 2005:514), it was surprising to discover that this correspondence may have a **PAsn** provenance (↓VGR?) or may result from areal convergence (↔LLR?). For example, *h* to *k* is found in a number of **PTB** allofams. The following examples are from Matisoff (2003:593): **PTB** **ha~*ka* ‘earth/ground/soil’, **PTB** **ha(:)k~*kak* ‘gag/choke’, **PTB** **hi:l~*ki:l* ‘bind/twist/roll/angle’, and **PTB** **hu~*kəw* ‘steal’.⁷ Note that this correspondence is also found in MK reconstructions: **Proto-Monic** **kntaam* ‘(fresh-water) crab’ > **PNyK** **kəntaam*, **PMon** **hətam* (Diffloth 1984:77.N52), **Proto-Monic** **kŋciəm* ‘bird’ > **PNyK** **kəŋciəm*, **PMon** **həcəm* (Diffloth 1984:71.N30), and **Proto-Monic** **gnaay* ‘to chew [V.tr]’ > **PNyK** **khəŋaay*, **PMon** **ŋŋaɨ* (Diffloth 1984:249.V268).

PAn/PMP **aa* to **PST/PTB** **aa* appears to be the most stable possible vowel correspondence. **PAn/PMP** **u* to **PST/PTB** **u* and **PAn/PMP** **ə* to **PST/PTB** **a* also appear to be recurrent. Vowel correspondences require further investigation, however; additional data (more possible cognates) are needed. References on ablaut changes in PST/TB must be considered in determining vowel correspondences (*cf.* Matisoff 2003, Ch. 12).

Some lexical evidence demonstrates possible shared morphological prefixes between PAn and PTB—see numbers 2 (right side) and 4 (walk/go) in Appendix 1.1. Some monosyllabic PST/PTB words appear to correspond across two Pan syllables (e.g., 1, 2, 4, 6, 9, 12, 15);⁸ In other cases PTB words appear to correspond to one or the other PAn syllable. The fact that the PTB monosyllabic word may correspond variously to different PAn syllables may be helpful in attempting to reconstruct the stressed syllable in Pre-An or PAn (see Larish 1999:371-376, cited in Pittayawat Pittayaporn’s paper at 10-ICAL).

⁶ For a discussion on *r* and *y* variation in Burmese, see Benedict (1972:41, especially footnote 134).

⁷ Since I could not find the symbol that represents allofamy—overlapping “more than” (>) and “less than” (<) signs—I have used the tilda symbol (~) instead.

⁸ This phenomenon reminds me of my E.S.L. teaching days in Japan, where the beginning student of English might say [sinku] for “think,” a disyllabic form corresponding to a monosyllable.

Unfortunately, I still have not been able to obtain a copy of Sagart, Blench, Sanchez-Mazas (2005). As far as I understand Sagart's position, he has been working toward establishing a genetic connection between PAN and PST for a number of years now. For Sagart to posit that PAN and PST comparanda are cognate, he expects that both the affix and root agree (pers. comm.). When considering the whole range of Austro-Asian languages, however, I suspect that a consistent correspondence between both affixes and roots will not be found due to the large number of languages that are monosyllabic. Thus, if recurrent **PASn** sound correspondences can be established, efforts must be made to reconstruct a corpus of **PASn** roots without regard to affixes. Core roots can be found that reflect genetically shared forms when affixes are peeled off. If **PASn** began as a monosyllabic language, then the development of prefixes, infixes, and suffixes may have been a secondary phenomenon in certain branches of **PASn** and may ultimately provide evidence for subgrouping. Since the evolution of languages often parallels biological evolution, it seems likely that a monosyllabic to polysyllabic development may better explain the evolution of **PASn**, just as complex plants and animals evolved from single-celled organisms.⁹

In some cases, one independent PTB monosyllabic lexeme may correspond to the penultimate syllable of the PAN word and another distinct monosyllabic lexeme with the same or similar meaning may correspond to the PAN ultimate syllable. Here are two examples taken from Appendix 1 that illustrate this phenomenon:

30 & 31	Proto-Forms	Glosses	Sources
PAN (30)	*qasiRa	salt	125A/200 Blust (97)
PTB	*la	salt	JAM 03:599L, 665R
	*m-t(s)i	salt/yeast	JAM 03:617L, 665R
	*tsa	salt	JAM 03:616C, 665R
PAN (31)	*timus ^a	salt	125B/200 Blust (97)
PTB	*m-t(s)i	salt/yeast	JAM 03:617L, 665R
	*tsa	salt	JAM 03:616C, 665R

The data in 30 and 31 suggest that the PAN word may have developed by compounding monosyllabic words with similar meanings.

51 & 52	Proto-Forms	Glosses	Sources
--------------------	-------------	---------	---------

⁹ Grace Odal-Devara (pers. comm.) informed me that E. Arsenio Manuel, an Anthropologist in the Philippines, has hypothesized that Proto-Asian may have been originally monosyllabic (Documenting Philippine Asian, 1996?).

PAn (51)	*esa	one	197A/200 Blust (97)
PTB	*t(y)ak	one/only	JAM 03:616C, 660R
PAn (52)	*isa	one	197B/200 Blust (97)
PTB	*ʔit	one	JAM 03:583C, 660R

This phenomenon suggests that PAsn and early Pre-An may have been originally monosyllabic languages. Proto-Asians may have begun with a largely monosyllabic lexicon and through a process of linguistic accretion—compounding, prefixation, suffixation, infixation, and *especially* reduplication—Pre-An may have begun to develop its trend toward largely disyllabic forms. Perhaps, Tai, OC, and PTB conserve original features of PAsn. Consider the ubiquitous sentence-final particle in Austro-Asian languages. Although some rare cases of disyllabic S-final particles can be found, the majority are monosyllabic. As a speaker of three Austro-Asian languages (Moklen, Thai, and Japanese), my impression is that there is a high degree of phonosemantic similarity in the form and function of S-final particle among these three languages. Do they represent archaic fossilized monosyllabic elements from a distant Asian past? For now, let me present examples from Mandarin and ask my Austronesian colleagues to e-mail me if they see parallels in the languages that they study. Li and Thompson (1981, Chapter 7) present numerous examples and extended discussion on the following six S-final particles in Mandarin Chinese: *le* ‘Currently Relevant State’, *ne* ‘Response to Expectation’ [*ne/ne:* in Japanese and *na* in Central Thai], *ba* ‘Solicit Agreement’, *ou* ‘Friendly Warning’, *a/ya* ‘Reduce Forcefulness’, and *ma* ‘Question’.

A possible problem with Benedict’s (1975, 1990) PAT and PAJ reconstructions is that he often takes the PAn form as conservative. That is, he reconstructs a PAT or PAJ form similar to PAN and then posits erosion to get monosyllabic languages (see Benedict 1990:20).

2.2.1. Austronesian and Sino-Tibetan/Tibeto-Burman Affinities

Taken as a whole, the possible cognates presented in Appendix 1 and 2 generally support the **PASN-Hyp**. As far as individual comparanda go, I am uncertain whether future research will support each case. However, in sixty-two data sets in Appendix 1, we find examples of PAn/PMP etyma that may be cognate with PST/PTB forms, possibly reflecting PAsn roots.

Sino-Tibetan is one essential “missing link” in establishment of a genetic connection between Austro-Asian languages. The data presented in Appendix 1 tentatively suggest that this link may be eventually connected. ((I have much more data to include in Appendix 2 but could not due to time

constraints.)) For the time being, I propose that establishment of recurrent PAsn sound correspondence must be left to future research.

2.2.2. Austronesian and Austroasiatic Affinities

Three potential arguments can explain the affinities between Austronesian and Austroasiatic languages:

- (i) A vertical genetic relationship (VGR) between **Pre-An** (equals Austro-Tai?) and Austroasiatic (Munda and MK)
- (ii) Remote contact between **Pre-An** and **PAA**, i.e., Old Lateral Loan Relationships (see Figure 1)
- (iii) More recent contact, i.e., Recent Lateral Loan Relationships

The first possibility (i), usually referred to as the ‘Austic hypothesis’, is too chronologically remote—Diffloth (1994) tentatively suggests ten to fifteen thousand years B.P.—to account for some lexical affinities between mainland-AN and MK languages. Relative to the extreme time depth required by the Austic hypothesis, more recent contact (possibility iii) is suggested when words group in a fairly limited set of semantic domains (Larish 1991:8). The Austic controversy (i) has been percolating for over ninety years since Schmidt (1906) coined the term ‘Austic’. For articles in support of Austic, see section 1.5.

2.2.2.1. PMM and PChm Long Vowels: Austic Retention?

The affinities between Austronesian and MK are due either to prolonged contact or a genetic relationship. If genetic, this would imply that Proto-Austic had a vowel-length distinction like Proto-MK. Cowan (1991:2) argues that if ‘the feature of vowel length distinction, which is a grammatical phenomenon, has a common genetic MK/PAN origin . . . it must have been preserved in PCA [= PAC] through PAN . . . confirm[ing] that those who had concluded already long ago that PAN originally must have had vowel length distinction were right’.

Table 1. Reconstructed Austronesian and Mon-Khmer Vowel Length Compared

	<i>dog</i>	<i>fish</i>	<i>wood</i>	<i>left (side)</i>	<i>bone</i>
PMK	*ʔa-c(ʉə)ʔ	*ʔa-kəʔ	*kəjh(u:)?	*jwi:ʔ	*j-l-ʔa:ŋ
P-Monic	*clur	*kaaʔ	*chuuʔ	*jwiiʔ	*j[l]uut
PNyK	*[khl/ch]ur	*kaaʔ	*chuuʔ	*chəwjiʔ	*chəlʉut
PMon	*klə	*kəʔ	*chuʔ	—	*cʉt
PMACM	*ʔasu[:](ʔ)	*(?i)ka[:]n	?*kayi:w	*ka-uiri[:]	*tuqəlá[:]ŋ
PMM	*ʔəy	*(?ɛ)ka:n	*kaʔi:[ow]	*kələy	*kəla:n

					*kɔla:n
PAC	*ʔase̞i *ʔase̞i??	—	*kayu:	—	*tulá:ŋ
<Ach>	<ʔaseə>	<uŋkot> <uŋkət>	<kayəə>	<wiə>	<tuluwəŋ>
PChm	*ʔasəu	*ʔika:n	*kaiəu	*ʔiãu	*tula:ŋ
PMal	*asuʔ	*ikan	*kayuʔ	*kA-iri *kibaʔ	*tulaŋ
PBtk	—	*dəŋkəy	*kayu	—	—
PPh	*ʔá:su	—	*ká:yuh	*kǎ-wĩyíʔ	*tuqǐlǎŋ
PMP	*ásu	*iSəkan	*kaSiw	*ka-wiyi	*tuqəlǎŋ

The data in Table 1 demonstrate a strong correlation between **PNyK**, **PChm**, and **PMM** long vowels, on the one hand, and **PMon** and **PMal** short vowels on the other. This correlation suggests that there may have been two separate linguistic areas influencing early Southeast Asian history and/or divergent shifts in stress placement (see Larish 1999, §7.1.2). The data show that the branches of **PMACM** have been subject to dissimilar areal influences. Although the boundaries and time depths of these areas are beyond the scope of the present study, they provide interesting possibilities for further research.

2.2.2.2. Lexical Evidence for Austric or Remote An-AA Contact

In addition to the core vocabulary examined in Table 1, a considerable number of additional lexemes, mainly anatomical terms, demonstrate strong An-AA affinities. Diffloth (1994:30) suggests a ‘probable Austric’ connection in the word for ‘tongue’. Comparing **PMP** **dilaq* ‘tongue’ with **Proto-Monic** **klntaak* ‘tongue’ shows that **PMP** *-*q* corresponds with Proto-Monic *-*k*. It is easier to account for **PMM** **kəlǎ:ʔ* ‘tongue’ as a reflex of Austric (or an MK loan) than it is to compare this etymon with **PACM** **dilah* ‘tongue’. To do so, we must posit a number of sound changes (e.g., **PMM** **kəlǎ:ʔ* ‘tongue’ < **Pre-MM** **təlǎ:ʔ* < **Pre-MM** **dələi:ʔ* [**d* → **t* → *k/_lʔ*]). Since Pearic (Pr) dialects show an alternation between *t* and *k* in *C_f* position for ‘tongue’ (Pr *rətət*, *srətək*, *kəltət*, *səltək*, *kətək*, *tək*), where **PMP** has **q*, this feasible correspondence supports the possibility that **PMACM** had a *t/k* alternation medially in the words ‘armpit’ and ‘knee’ (first compared in Larish 1999, Table 7.9). When **PMM** **puʔa:k* ‘armpit’, **PChm** **paʔa:ʔ* ‘armpit’, Mal *kətiak* ‘armpit’ (cf. **Proto-Monic** **knlak* ‘armpit’) and **PMM** **tuʔət* ‘knee’, **PChm** **(t/kǐ)ʔut* ‘knee’, Mal *lutut* ‘knee’ are compared, the ultimate syllables of Malay and the mainland-AN languages appear to show good correspondence, but the penultimate syllables do not. Since Malay has -*t*- in both ‘armpit’ and ‘knee’, this suggests that in some cases where *-ʔ- is reconstructable for **PMACM**, there may have been an alternation between *t* and *k* in medial

position. If so, the medial glottal stop could have been conditioned by dissimilation of the medial consonant before final *t* or *k*. According to Diffloth (1984:97.N117), Mon dialects normally use a compound form **pāñ(kə)nək* for ‘armpit’, which shows a strong affinity to **PMM** **puʔa:k* ‘armpit’ and **PChm** **paʔa:ʔ* ‘armpit’. The close correlation between these forms suggests that they were borrowed from a Monic language nearly at the same time and/or place (or that **PMM** borrowed from **PChm**), but the addition of data from Borneo (cf. Biaju Dayak pènang; Maanjan apènang ‘upper arm’) may suggest an Austric connection.

The evidence for ‘heel (of the foot)’ is noteworthy considering that Moklen and Chamic words look related to MK on the one hand, and Moken and Acehnese appear cognate on the other (cf. Mkn-Dng *kʰənĩ:m* ‘heel’, Ach-Lws *geunòë*).¹⁰ The Moken and Acehnese forms correspond phoneme-by-phoneme, except for the unexplained final *-m* of Moken, which could be a suffix. Compare Mkl-BDC *kadurŋ* and Mkl-Lmp *cadəŋ* with Rad *kduł*, Jor *kəʔdul*, Rog *kaʔduən*, Rad-DB *kʔdu:l* ‘talon’, Chm-Mou *kadol-takay* ‘talon’, Rad-DB *jəŋ* ‘pied, patte, jambe (foot, paw, leg)’ (< **PChm** **kaʔdu:[nl]* < **Pre-Chm** **kaldu:[nl]*) ‘heel’. The Moklen and Chamic forms correspond across both syllables, whereas only the major syllables of An and MK forms appear cognate: **Proto-Monic** **juŋ*, **PNyK** **cʊŋ*, **PMon** **caʔŋ* ‘foot’ (Diffloth 1984:97.N120).¹¹

Another significant anatomical term is ‘neck’. There is a high degree of lexical overlap between Austronesian, Mon-Khmer, and Thai words in the block of data presented below which is complicated by the fact that it is difficult to establish the exact meaning of a lexeme by simple one-word glosses.¹² A tentative hypothesis is that forms with *t*- initials mean ‘nape’, those with *l*- initials mean ‘front of the neck’, and *b*- initial forms mean ‘throat’. This is supported by ItgB *ʔliyig* ‘front (of neck)’, ItgB *tʔngid* ‘nape’, where the distinction between the front and back of the neck is clearly evident (Reid 1971.203.192).

In reconstructing **PMACM** **tʔ(ŋ)qu:[ŋkʔ]* ‘nape of neck’, MAC languages do not retain the prenasalized consonant seen in Mal *təŋkok*, but **PAC** and **PMM** agree in having heavy major syllables. In addition, the mainland languages give us a range of final velar consonants to examine. As Thai and MK languages show strong affinities to **PMACM**, this word may

¹⁰ Lewis (1960:66) compares Mkn-Lws *kenim* ‘heel’ and Ach *geunòë* ‘id’.

¹¹ Skeat and Blagden (1966:630.H69-70) first observe curious similarities between Chamic, Aslian, Nicobarese, and Andamanese words for *heel* (e.g., cf. Andamanese Beada *ta:rkʊ:du:l-de* and Kol *o:m-ke:tel-che*).

¹² The word *neck* involves a number of related semantic domains (e.g., nape, throat, base of the neck, top of the neck, Adam’s apple, and so on).

have Austro-Thai and/or Austric connections (cf. **Proto-Monic** *kɔɔʔ, Thai-C *kʰɔː*).

neck (nape) 25/200, Mkb84, Benj72, Ray123, Reid192

PMACM	*tiku[:]ŋ	‘nape of neck’
PMM	*(tʷ)kɔːŋ/*[n/l]ʷkɔːŋ	‘nape of neck’
Pre-Mkn	tukʰɔː	‘nape of neck’
Mkn-Dng	tukʰɔ	‘back of neck’
Mkn-VP	tukʰɔː	‘neck’
Mkn-Lws	tukoʔ	‘neck’
Mkn-Rw	tukʰoː	‘the neck’
Pre-Mkn	*[n/l]ʷkɔːŋ	‘neck’
Mkl-BN	nukɔːŋ/(lu)kɔːŋ	
Mkl-BDC	ʷkɔːŋ	‘neck’
Mkl-Lmp	lakɔːŋ	‘the neck’
Mkl-Lmp	gəhɔʔ	‘to be necked (of a nose)’
PAC	?*takuəi	
Ach-Col	takoeë	
Ach-Saw	lihiə/lihe	‘neck’
Ach-SkBl	takuə	
PChm	*takuai	‘neck’ > Rad kkuε, Jor təkuaɪ, Rog takuai, Chm təkoy [Chm-Mou takɔy (phɔːʔ) ‘nuque (nape)’, Jar-SkBl təkoi, Chm-SkBl takuəi]
PMal	*lihər > Mal leher, UL lihel; Mal təŋkok	‘nape of neck’
Mal	anak təkak	‘uvula’
PMP	*liqəy	‘neck’
Png	bíkliw	
PPh	*lí:qəy	‘neck’
[PhML]	Agta lig, Atta bullaw, Blw ba'gang, Btk liqig, BilK liqal, BilS liqal, Bkd 'liqig, BonG ba'gang, DgtC 'leg, Gad buqlaw, IfgAg ba'gang, IfgBt bagang, IfgBy bagang, Isg buqlaw ‘neck, throat’, IvtI ragaw, ItgB 'liyig ‘front (of neck)’, ItgB ti'ngid ‘nape’, IvtI lagaw, Klg liyig, KlaG ba'gang, KlnKy bukliw, KlnKl buklew, KnkN ba'gang, Mmn liqig, Ata liog ‘throat, neck’, MbAD 'liqig, MbI liqig, MbKC liqig, MbS liqʌg, MbT liqig, MbCW liqig, Msk liyig, Sml killong, SblBt 'liyiy, Snl 'rerik, San leheq, SubS dliig, SubSc glig, Tbl lihəl, TbwA liqig, TbwK dikil, Tsg liqig	‘neck’
Tag	batok	‘nape’
Tag	liig/leég	‘neck’

UAN	*lihiy	‘neck’	
Thai-C	k ^h ɔː	‘neck’	
[Mon-Khmer]	Proto-Monic *kɔʔ, PNyK *kɔʔ, PMon *kɔʔ		
BM	kɔʔ, TM kɔʔ, NyK kɔːʔ	‘neck’	
Khm	kɔː	‘neck’	
Nic dialects	oŋ-lô·ŋa, ŋeː, uain, en-lô·ŋa, aŋ-lô·ŋa, lik-kun		
Pr	kɔːk	‘neck’	
Asl-Benj:	Ks təŋkɔŋ, Je təŋkɔŋ, BN təŋkɔk, Sn təŋkɔŋ, Sa təŋkɔŋ, LY təŋkɔŋ, KB cəŋɔd, Mr ʔuŋut, BD nut, Mt ʔəŋəut, LJ taŋən, Tm taŋən, Sml taŋən, SmlI taŋən, CW ləŋeʔ, JH ləŋeʔ, Sl ləŋeʔ, Tq ləŋeʔ, SB səmaroŋ	‘neck’	
Sen	lɔŋi	‘neck’	
Sen	taŋkok	‘nape of neck’	(MLW?)

neck (throat) Reid315 (throat)

Mkn-Dng	bulɔːŋ	‘throat’	
Mkn-KS	buloːŋ	‘neck, throat’	
Mkn-Lws	bulong	‘neck, throat’	
Mkn-Rw	buloːŋ	‘the neck, the throat’	
Mal	kəroŋkoŋan	‘throat’	

Other lexemes that support the Austric hypothesis are briefly considered here. Another case involves **PMACM** *ri[ʔ]buŋ ‘bamboo shoot’ and **Proto-Monic** *tɔaŋ ‘bamboo-shoot’ where the major syllables show a partial phonological resemblance, yet **PPh** *(qa)Dəbuŋ shows a broader resemblance across two syllables.

Since BonG *ta'mong*, Ata *qapongag* ‘cheek’ (Reid 1971:59.45) are likely cognate with Mkn-Dng *təbɔːŋ* ‘cheeks’, Mkn-Rw *təboːŋ* ‘the cheeks’, Mkl-BDC *kəbɔːŋ* (fast)/ *kəwɔːŋ* (slow), and Mkl-KY *tabɔːŋ*, affinities with Aslian could go back to Austric, albeit Aslian could have borrowed from an Austronesian source (cf. Sem Pa. Max, Sem Buk. Max *kəbang*; Sem. K. Ken. *kabaː* or *kapá*; Tembi *kapong* or *kapang*, Sem. Plus *peng-peng*; Sen. Cl. *məng* (Skeat and Blagden 1966:556.C81,84). Note also that Sen. Cl. *məng* appears connected with Chamic: Rad *mieŋ*, Jor *měŋ*, Rog *miaŋ* [Rad-DB *mieŋ* ‘joue’, Chm-Mou *mieŋ* ‘joue’, Chm-SkBl *amiöŋg*] (< **PChm** **miəŋ* ‘cheek, jaw’). Since the AN-MK affinities appear in a core-type word related to the body, this supports the Austric hypothesis. The m/b variants in C_m position and vowel differences in V_m are curious.

Another core-type word with Mkn-Mkl and Aslian affinities is ‘chin’. Compare Sem. Kedah *'ngkek* or *'ngkeː* ‘chin’ (Skeat and Blagden

1966:559.C113) with Mkn-KS *təkě:k*, Mkn-Lws *tekék*, Mkn-Rw *təki:k*, Mkn-Rw *təkĩ:k*, Mkl-BDC *tukw:k*, Mkl-KY *tukə:k*, Mkl-Lmp *təkə:k* (< **PMM** ^{?x}*təkě:k* ‘chin’). The fact that only the major syllables show phonetic agreement makes the possibility of borrowing less likely.

Two further items of core vocabulary support the Austric hypothesis:

PMACM *(mə-)nɪpi: ‘to dream’ > **PMM** *mɪpuɪ/*n/lɪpuɪ, **PAC** *lu(m)pi: > Ach-Dur lumpəə, **PChm** *lupɛi (possibly *lipɛi), **PMal** *impi > Mal bər-mimpi, UL mipi (< **PMP** *(mi)-Səpi, *nɪpi/*nupi); **Proto-Monic** *([k]m-)pɔɔʔ ‘to dream’ > **PNyK** *([]m-)pɔɔʔ, **PMon** *kəpɔʔ > Pr phok/phoʔ, Khm yùəl sɔp(t), Nic enfu:a-chakâ, Asl-SkBl Sem. Kedah ěm-pa’, Lebir ěm-pa’, Kerbat (ěm-pak), Krau Tem. ěm-púa’, Sen. Cliff. ěm-po’;

PMACM ^{?x}*kikít ‘to bite’ > **PMM** *mɔkɔ:t, *n/lɔkɔ:t, **PChm** *keʔ (possibly *kɛʔ), **PMal** *gigit > Mal gigit, UL gigeʔ (< **PMP** *katkat/*kətkət/*kitkit/*kutkut (poss. unified as **Pre-MP** ^{?x}*kitkít); **Proto-Monic** *kɯt, **PNyK** *kɯt, **PMon** *kɪt > BM kic, TM kic, NyK kit; Pr khat, Khm cɤk (cf. Thai-C kàt),¹³

2.2.2.3. Morphological Evidence for Austric or Remote An-AA Contact

Morphological evidence for a genetic connection between Austronesian and Austroasiatic has been examined by a number of scholars (cf. Reid 2005, 1994). Such evidence supplies considerable weight to the Austric hypothesis, given the fact that morphological systems are not likely to be borrowed.¹⁴ For example, Nils M. Holmer, a megalocomparativist (according to Matisoff 1990), notes the striking resemblance between prefixation and infixation throughout Asia:

In Austronesian the ‘derivational’ prefixes are based chiefly on the consonants *p-*, *t-*, *k-*, *m-*, *s-*, and zero (which latter might be identified with one of Dyen’s ‘laryngeals’) . . . The interesting thing is that the same type of consonant reappears in written Tibetan, e.g., the *b-*, *d-*, and *g-* of the verbal inflexion, further the initial *’-* or *h-*. In Tibetan words beginning with consonant clusters we often notice *b-*, *d-*, *g-*, *m-*, *s-* as initial consonants. The same is the case in reconstructed Ancient Chinese, where the prefixes take the forms **p-*, **t-*, **k-*,

¹³ In addition to this set ending in final stops, see ‘bite, v.’ in Appendix C for numerous instances of bilabial finals (e.g., Ach *kap*, Nic *opkâp*, Asl *kap/kab*).

¹⁴ Thieme (1964:589) states that ‘morphological correspondence is . . . the safest indication of genealogical relationship’.

**m-*, **s-*, and **h-*.¹⁵ Among the ‘derivational’ formatives we also have to consider the infixes, chiefly with *-n-*, *-m-*, *-l-*, and *-r-*, which play an important part in the Austronesian languages. They are also found in Mon-Khmer (*-l-*, *-r-*, *-n-*, and *-m-*) and in Munda (*-l-*, *-r-*, *-n-*) (Holmer 1963:21).¹⁶

Furthermore, Matisoff (2003:88) suggests that the relative age (ancient vs. recent) of prefixed elements or “formatives” in **PTB** correlate with semantic opaqueness (ancient prefixes) and semantic transparency (“relatively recent”); he discusses seven **PTB** prefixes: **s-*, **m-*, **a-* [$\int(\ominus)$], **b-*, **g-*, **d-*, and **r-* (2003, 4.2 to 4.4), noting that the first three (**s-*, **m-*, **a-*) are ‘highly important, with relatively well-defined semantic content’ (2003:87-8).

The close resemblance between Moken-Moklen and Aslian affixes warrants repetition of the following quote from Skeat and Blagden’s ‘Grammatical Notes’, augmenting Holmer’s summary:

The common verbal prefixes are Sēmang *ma-*, Sakai *ǎm-* (*ěm-*), *ně-*, *ně+a* consonant anticipating the final consonant of the root (this last seems to be typical of Northern Sakai, but appears also to occur in Sēmang, and may be of Sēmang origin) *hi-*, *ki-*, and Běsisi *ka-*. *Pa-* (with variants *pi-*, *pě-*, etc.) and *ta-* (*tan-*, *tě-*, *teng-*) in several dialects form causal and sometimes transitive verbs. In Běsisi *na-* and *ta-* form adverbial and adverbial demonstratives. There appear to be many other prefixes.

The commonest infixes are *-ěm-* (*-am-*, *-um-*) and *-in-* (*-ěn-*, *-an-*). It is curious that these (like some of the prefixes) are common to the Mon-Khmer and Malayo-Polynesian families, still more curious that Sēmang and Northern Sakai frequently use the *-in-* infix in words with Malayan affinity, though it is as good as non-existent in Malay itself.

The Jakun dialects use Malayan prefixes and suffixes, not, however, always absolutely identical with their ordinary Malay forms: *meng-* is sometimes represented by *ma-*, *běr-* by *ba-*, *di-* is not necessarily passive in Jakun (Skeat and Blagden 1966/06:774).

To qualify as evidence for common heritage, cognates between MK and Austronesian languages must be found in Austronesian languages distant from mainland Southeast Asia. The diffusion of MK lexemes might extend as far as the Philippines (e.g., alcohol, eggplant).

The criteria for distinguishing remote contact or common origin from more recent diffusional resemblances must be considered. In more remote cases, phonetic agreement might only be found in major ultimate syllables. Larish (1991) considers such cases for the following reasons:

15 Holmer also notes the parallel between these initial consonants and those of Old Mon.

16 Holmer neglects to provide examples from Sino-Tibetan. Wolfenden (1929) notes the *-r-*, *-l-*, and *-s-* infixes of Tibetan, *-ri-* in Dima-Sa (Bodo), *-r-* in Mikir (Naga-Bodo) and *-rr-*, *-r-*, *-ra-*, and *-ro-* infixes of Old Kuki and their identity with Mikir *-r-*.

First, MAC [Moken, Moklen, Acehnese, and Chamic] and MK languages erode phonologically on the left; consequently, one might expect main syllables to be relatively more stable over time than presyllables and minor syllables. Second, MAC and MK languages optionally drop non-ultimate syllables (Larish 1991:7).

Recent contact is suggested whenever lexical affinities are shared only among MK and mainland-AN languages; recent loans usually demonstrate phonetic similarity in both major and minor syllables.

Headley (1976b) and Diffloth (1994) are useful sources, for they consider the directionality of possible and probable loans between Austronesian and Mon-Khmer languages. Out of 96 comparanda, Headley (1977b) proposes that 72 sets are ‘words of probable Mon-Khmer or Austroasiatic origin (pp. 454-64), 11 sets are ‘words of probable Indo-European origin’ (pp. 464-6), 9 sets are ‘words of uncertain origin’ (pp. 467-8), yet only 4 sets are identified as ‘words of probable Austronesian origin’ (p. 466). Given the probable importance of Austronesian groups in the early Southeast Asian history, as argued in Larish (1999, Chapter 8), we should expect more instances of Austronesian loans being incorporated into MK languages (e.g., spoon, swim, needle). Although Headley (1976b:469) only briefly addresses the Austric question, he presents fourteen items of basic vocabulary between AN and MK (*ibid.*:470). Again, I do not support a genetic connection in all the cases that Headley identifies as basic vocabulary: back, bird, catch, cut, to fly, intestines, leaf, leg/foot, pull, river, sand, swim, warm, and wash. Instead, the data indicates mixed etymologies, both genetic (↓VGR) and diffusional (↔LLR). The data in Table 2 suggest a possible Old Lateral Loan Relationship (↔OLLR) between **PMM** and **PAC**:

Table 2. Evidence for Contact between Acehnese-Chamic and Moken-Moklen (excluding Malay)

gloss	MK	PMM	Ach	PChm	PMal
bird	Nic-Car checho:n	^x cicum/ ^x (ti)cum	<cicem>	^x cim	*buruŋ
nephew	PMonic *k()muun	^{?x} kamə[:]n	—	^{?x} ʔamuə̃n ^{?x} kamuə̃n	—
fruit/egg (clsf)	Asl-Sn <bəh> ‘fruit’	Mkl-BDC <p ^h əh> ‘classifier’	<bəh>	*boh ‘egg, fruit, ball’	*buah
fish hook	NyK <chwək> ‘hook, v.’	*c ^h əwək *c ^h /suwiək	culawīə? (with infix)	Chm-Mou <wəh> ‘hameçon (fishhook)	mata kail

The fact that Moken-Moklen, Acehnese, and Chamic (MAC) share a number of MK loans and parallel structural convergence toward MK, where Malayic languages apparently do not, supports that the MAC languages may have been in contact after Acehnese-Chamic broke off from Malayic. It is probable that the MK influence first began when the MAC languages were in contact due to the strong structural parallels exhibited by the mainland-AN languages. It seems unlikely that these words could have been borrowed at separate times and places. It is easier to posit one case of borrowing rather than two or three. Therefore, the data in Table 2 support Cowan's (1948) hypothesis for a special (*betrekking*) relationship between MAC & MK (see Larish 1999, Book 1, p. 57 for a summary).

The evidence for 'bird' deserves detailed examination since MAC & MK comparisons illustrate a full range of possibilities, including sound symbolism and chance similarity. These last two possibilities are rarely an issue. The mainland-AN and MK languages share such an abundance of structural and lexical affinities that chance similarity is ruled out by the quantity and quality of the evidence (*cf.* Shorto 1975, Larish 1991, Larish 1997). In comparing **PMM** **(ti)cum* and Ibl, KlnKy (Phil) *titit* 'bird' (Reid 1971:51), however, sound symbolism becomes an issue, as both forms could be onomatopoeic or imitative. Even if they are, by comparing their rhymes, it becomes readily apparent that **PMM** **(ti)cum*, Ach *cicem*, and **PChm** **cĭm* are more similar to MK (e.g., Nic *shichu:ə* compared with Vtn *chim*) than to the Philippine forms. Sound symbolism may account for the affinities between the Southeast Asian mainland languages (including MK) and Ibl, KlnKy *titit*. In cases where such an explanation is not available, affinities between MK and insular-AN languages must be considered possible evidence in support of (i), the Austric hypothesis. When **PChm** **cĭm* with Ach-Col *tjitjém*, Ach-Dur *cicem* 'bird' are compared, it appears that the Chamic languages dropped the presyllable. If so, **PAC** ^{?x}*(ci)cim* can be reconstructed. Comparing **PMM** ^x*cicum*/^x*(ti)cum* and **PAC** ^{?x}*(ci)cim* suggests **PMACM** ^{?x}*(ci)cim*, where *i* → *u/_m* and *i* → *i* through vowel harmony. Another possibility, that of separate MK influence, also appears feasible. The *C_p* and *C_m* consonants of Mkn-KS *cicum* and Ach *cicem* appear closest to Nic-Car *checho:m*, and the initial consonant of the most commonly heard Moklen form *ticum* may correspond to Proto-Monic **k-* in **kɲciəm*. Rad, Jor *cĭm* 'bird' and Chm *cim* 'oiseau' (< **PChm** **cĭm*) are identical with Aslian Sn, Sa, MM, Sl *cim*, yet similar to Vtn *chim*. Rog *cip* is closest to Aslian LJ, LY, Tm *cɛp* and Tq *cip*. The fact that Chamic languages share the same *C_f* alternation with Asl-C and Asl-S suggest a possible contact relationship. Unique lexemes such as Mal *buruŋ* 'bird' and Asl-Ks *kawaw* 'bird', which is representative of Asl-N, obviously show no connection to the data which strongly supports the EMAC-MK *sprachbund*, first discussed in Larish (1999, §7.2).

2.2.2.4. Kinship Terms

In discussing the kinship terms in Larish (1999, Table 6.9), the AN-MK affinities are attributed to borrowing, hence they are labeled **PMACM** lexical innovations. Cowan (1991:10-11), however, suggests an alternative hypothesis. He claims that the similarities between the mainland-AN and MK words for ‘grandchild’ are Austric retentions in Acehnese and Cham. If such is the case, then **PMM** *[t/c]ɔcǝʔ ‘grandchild’ could also be an Austric retention. An important factor is that the mainland-AN languages have been in contact with MK speakers much longer—perhaps two millennia as claimed by Thurgood (1999)—than most insular-An languages. Two thousand years ago, early Austronesians and Mon-Khmers who were in contact probably recognized their linguistic kinship to a better degree than we can today. This being the case, they may have tended to retain mutual linguistic elements that aided in communication, elements that would be archaic in other areas of Austronesia.

We should note one more lexeme that was probably borrowed from MK by speakers of **Pre-Mkl**, one which is not evident in Moken, Acehnese, or Chamic:

Pre-Mkl **bulaw* ‘wife’ > Mkl-BS *bilaw*, Mkl-BDC *bulaw*, Mkl-KY *bulaw*, Mkl-Lmp *bəlaw* ‘a wife’;

Proto-Monic **braw*, **PNyK** **phrəw* ‘wife’, **PMon** **prɛə* ‘woman, wife’ (Diffloth 1984:115.N174), BM *hayaŋ-prəa*, TM *prəa*, NyK *phrəw* ‘wife’, BM *maki*, TM *nih-prəa/maki* (old), NyK *phrəw-phrəw* ‘woman’, Pr *čhəŋi(:)n*, Khm *prəpə̀an(th)*, Nic *kān/kāna*, Asl-SkBl Ben. New. *malau* ‘woman’, *malaulau* ‘wife’ (Skeat and Blagden 1966:601.69)

Clearly, **Pre-Mkl** **bulaw* ‘wife’ shows a closer affinity to either NyK or Aslian here, but it matches closely with **Proto-Monic** **braw* ‘wife’, suggesting that it may be an old loan from MK.

2.2.2.5. Animals

Except for Nicobarese and Andamanese, an important distinction must be made between land-based Mon-Khmers and island, coastal, and river-based Austronesians. This opposition allows us to predict somewhat the direction of borrowing. By and large, most of the lexemes in the plant and animal domains appear to be borrowed from MK into An, especially for land-based plants and animals. The large number of apparent loans in the animal domain that were borrowed from MK languages into Moken-Moklen, Acehnese, Chamic, and Malayic suggests that these early Austronesian speakers settled on the mainland after the MK, adopting many of their names.

The following list is representative. The reader is referred to the data in Larish (1999, Book 2, Appendix C):

ant
 bat; butterfly
 bird
 bird-of-prey (hawk, eagle)
 crab (generic)
 crow, n.
 dog
 dove; pigeon (see lexeme 1)
 duck, n.
 fly (insect) (see lexeme 2)
 fly, v.
 nest, web
 spider
 squirrel
 tiger
 wing

The following extended examples are typical of the type of lexical similarity found between MK and mainland-AN languages. In this first case, borrowing is suggested by the close phonetic agreement between MK and An etyma:

PMM ^x?ada: ‘duck, n.’ (prob. MK loan)
 Mkn-Dng ?əra ‘duck’
 Mkn-KS ?ada: ‘duck’
 Mkn-Rw ?ada: ‘duck’
 Mkl-KY ?ada: [ʔa.²¹.’da:²³²]
 Mkl-Lmp ?a?da: ‘a duck’
 Ach-Saw ?ite?
 Chm-Mou ?atà: < ^x?ada: ‘canard’ (prob. MK loan)
 Mal itik, UL iti?
UAN *itik ‘duck’
 [Mon-Khmer] **Proto-Monic** *(?a)da:, **PNyK** *t̄a, **PMon** *(?/kə)t̄aɣ
 BM, TM ?atèa, NyK tàa,
 Vtn con v̄it
 Nic wet

The strong affinity between **PMM** ^x?a'da: and Chm-Mou ?atà: (< ^x?ada:) ‘canard’ suggests that **PMM** and **PChm** borrowed from the same MK source or, perhaps that speakers of **PMM** obtained the word indirectly from **PChm** or vice versa. Since most Aslian languages have borrowed *itek* from a Malayic language (Skeat and Blagden 1966, II:585; Means and Means 1987:124), neither Vietnamese *con vịt* nor Nicobarese *wet* can be the loan source.

The phonetic similarity between **PMM** ^x*cicum*/^x*(ti)cum kəla:ŋ* ‘bird of prey (hawk, eagle)’ and **PMon** **kəŋiəŋ* ‘kite’ is striking, yet in this case Austric retention cannot be ruled out. The possible correspondence between mainland-An *a:* and Monic *iə* is typical of a difference in tense/lax phonation type (Larish 1999, pp. 316-9, §6.2.3.2.2).

bird of prey (hawk, eagle) Lrb225, Hdly1.38

PMM ^x*cicum*/^x*(ti)cum kəla:ŋ* ‘bird of prey (hawk, eagle)’ (poss. MK loan)

Mkn-KS *cicum kəla:ŋ* ‘bird of prey’

Mkn-Rw *kəla:ŋ* ‘hawk’

Mkl-KY *ticum kla:ŋ* ‘bird of prey’

Mkl-Lmp *kla:ŋ ʔuk* ‘hawk’

PChm **kala:ŋ* revised as ^x*kala:ŋ* ‘hawk’

Thai-C *nók yǐ:aw* ‘bird of prey, hawk; eagle; kite; harrier’

Thai-S *nok*²¹ *yǐ:aw*⁴⁵³

[Mon-Khmer] **Proto-Monic** **liŋ-liəŋ*, **PNyK** **liŋ-liəŋ* ‘hawk: *Spizaetus sp.*’

PMon **kəŋiəŋ* ‘kite’

Nic *kalâng* ‘white-bellied sea-eagle (*Cuncuma leucogaster*)’

PCNB **klaŋ* ‘kite, eagle’

PJH **klaŋ* ‘kite, eagle’

PSBnr **klaŋ* ‘kite, eagle’

PTB **gləŋ* ‘eagle/vulture/falcon/bird of prey’ (Matisoff 2003:590) shows that this lexeme has a wide distribution.

2.2.2.6. Plants

The following lexemes are representative of lexical affinities between MK and An languages within the plant domain. The reader is referred to the data in Larish (1999, Book 2, Appendix C):

bamboo

bamboo shoot

bean

corn

fruit, n.; classifier

jackfruit

leaf

pandanus spp.

pepper (black)

pepper (red-hot)

root, n.

Examination of the plant lexicon suggests both archaic residue (↓VGR) and diffusional cumulation (↔LLR). Lexical relationships within the plant domain are characterized by different strata of historical connection, some possibly genetic (bamboo shoot, fruit), older loans (jackfruit), and others more recent loans (bamboo, bean, corn, papaya, pepper). Cases where a plant is not endemic to Southeast Asia are particularly important, for they can be easily identified as loans. For example, some plants (e.g. corn, papaya) were probably introduced into Southeast Asia subsequent to European exploration and concomitant discovery of exotic species in Africa, South and Central America, and other new areas. For example, in examining the comparanda for *corn* (endemic to South America) and *papaya* (endemic to Central America), we can see various lexical sets which suggest that the lexical diffusion of these words may have begun in different locations.

2.2.2.7. Aquatic Domain

Larish (1999, Chapter 8, Table 8.1) describes the early politico-cultural zones in Southeast Asia where Austronesians and Mon-Khmer speakers developed local trade networks. The fact that the Austronesians and Mon-Khmers may have originally lived in complementary environments is important in understanding why AN-MK interaction succeeded. If both groups had been in competition for the same resources, this would have led to conflict. Austronesians maintained a distinct advantage in ocean, coastal, and riverine environments, so in words related to aquatic domains—unlike plant and animal loans—we might expect that the normal direction of borrowing is from Austronesian into MK. The following lexical items are representative of this domain: boat, coral, crocodile, swim, v., wind (from the west), and monsoon.

2.2.2.8. Evidence for Early Contact between Moken-Moklen and Aslian Languages

A number of striking lexical and phonological affinities between Mkn-Mkl and Aslian languages are presented in Table 3. These lexical agreements appear to be most consistent between Mkn-Mkl and the southernmost representatives of the Northern Aslian subgroup, especially the *Mənrəʔ* (Mr), *Batəg Dəʔ* (BD), *Batəg Nəŋ* (BN), and *Ceʔ Wəŋ* (CW).¹⁷ In discussing the demographics of these Aslian tribes, Benjamin (1976:47) notes that *Mənrəʔ* live ‘along the mid-reaches of the Kelantan River around Kuala Krai and Bertam’ and also on the lower Lebir River; the *Batəg Dəʔ* live on the ‘Aring River in south Kelantan . .

¹⁷ For a map of illustrating the distribution of the three Aslian subgroups (Asl-N, Asl-C, Asl-S), refer to Benjamin (1976:46).

. ranging over into Trengganu and Pahang’; the *Bateg Nəŋ* dwell in ‘villages near Jerantut in Pahang’; and, finally, the *Ce? Wəŋ* live ‘on the southern slopes of Gunong Benom between Raub and Kuala Krau in Pahang’. Benjamin (1976:47) classifies the *Mənra?* and *Bateg Nəŋ* as ‘Negritos, the *Bateg De?* as “nominally ‘Negritos’, but with a large proportion of non-Negrito elements in their physical and cultural makeup,” and the *Ce? Wəŋ* as ‘non-Negritos’. Most of these Aslian people are ‘semi-settled’ in the locations described above, but some are nomadic. According to Benjamin (1976:47), the total population of these four groups is close to 800.

If speakers of **PMM** or **Pre-MM** came from the East Coast of the Malay Peninsula, as their descendants consistently claim (Larish 1993:1316, 1999:443), some linguistic trace of this movement should remain. Although limited in quantity, lexical and phonological agreements between Mkn-Mkl and the languages of southernmost tribes of the Northern Aslian subgroup are qualitatively striking enough to lend support to the oral history of the Mkn-Mkl. What is significant about these Aslian tribes is that they inhabit the rivers of northeast Malaysia. In terms of trying to deduce the former movements of the ancestors of the Moken-Moklen, the evidence in Table 3 points to a strong genetic (\uparrow VGR) or old diffusional relationship (\leftrightarrow OLLR) between the Moklenic and Aslian people. These affinities are presented in Table 3:

Table 3. Lexical Affinities between Moken and Moklen and Aslian Languages

	<i>flower</i>	<i>good</i>	<i>drink, v.</i>	<i>nurse, v.</i>	<i>small</i>
Asl-N	Ks, KB buŋa?, Je boŋe?, BD boŋa?, Mt boŋa?	Mr ?abən	Mr, BN, BD ?am	Sem. Buk. Max. ma:’ek am/ek ma:’am ‘to suckle’	Mr, BN, CW kanet
Asl-C	Tm boŋa:?, SmI boŋa:?, SmII boŋa?	SmI bə:r SmII bər	—	Sem. K. Ken. man- mä:m ‘teat’, Sen. Cliff. tē-u mem ‘to suck’	SB ke?net
Asl-S	MM boŋa?	—	—	—	Tq kanit
PMM	buŋa:?	*?amə:n *?amǒ:n	*(ma)?am	Mkl-Lmp <mam> ‘to suck milk	*nek (cf. Mkl kanek

				(of a baby)'	'small pot for cooking rice')
Ach	buŋɤŋ	---	<minɔm>	mom 'female breasts'	—
Chm-Mou	p̄iŋu:	<bon> ¹⁸ 'facile'	<muŋnum>	—	<neh> 'petit'
PChm	buŋã: poss. buŋi:	*siam	*munum	*mãm 'to suck'	Jar-Lws <net>
PMal	buŋa(?)	*baik	*inum	<məntéték> 'to suckle'	*kəcik *kəcil
PPh	bú:ŋah 'fruit'	*p̄i[y]a *ʔup̄i[y]a	*ʔinum	—	
PMP	*buŋa ?*buŋa?		*um-inum	UAN *dibdib 'suck (at the breast)'	*ke(n)Diq?
P-Monic	pkaaw	—	*sooŋ	—	—
PNyK	kaaw	—	*çooŋ	—	—
PMon	kaq̄	—	*sa̱iŋ	—	—

Other AN-Aslian data that show a *sprachbund* relationship (which excludes Malayic languages) include *breathe*, *true*, *root*, *drink*, *leaf*, *small*, *dog(?)*, *fruit*, *not*, and *dry*. The following data block from Larish (1999, Book 2, Appendix C) demonstrates affinities between Moklenic and Aslian:

bat; butterfly Lrb28, Lrb55, Ray26, Reid37 (butterfly)

Pre-MM ^{ʔx}k/tawaŋ, ^{ʔx}k/tiwiŋ, ^{ʔx}k/tib^wiŋ, or ^{ʔx}k/tawar

PMM ^{ʔx}kawan/^{ʔx}tawan 'bat; butterfly' (poss. MK loan)

Mkn-Dng kawan 'butterfly'

Mkn-KS kəbuŋ 'butterfly'

Mkn-Rw kawan 'a butterfly'

Mkl-BDC kawan/tawan 'bat, butterfly'

Mkl-KY tawan 'bat, butterfly' [ta.²¹. 'wan³⁴³] {ta:hwan²}

Mkl-KY tuwo:t 'bat'

Mkl-Lmp tawan 'butterfly'

¹⁸ Lewis (1960:47) suggests several affinities: Burm. *amuŋ*, Skt *bun*, and Pang Sam *a.bön* 'good'.

Mkl-Lmp	kawan	‘bat’	
Ach-Saw	baŋbaŋ/bambaŋ	‘butterfly’	
Chm-Mou	katit, ?inw: patà:y	‘papillon (butterfly)’	
Mal	kupu-kupu	‘butterfly’	
Mal	kəlawar	‘bat’	
PPh	*pǎní:ki	‘bat (fruit)’	
[PhML]	IvtI kulivaavang, Ivt kodibaabang, Klg kalibangbang, MbKC kilibangbang	‘butterfly’	
Tag	paruparo, mariposa (Spn loan)	‘butterfly’	
Tag	talibatab, paniki, bayakan, kabag-kabág	‘bat’	
UAN	*kalu’añ	‘bat’	
UAN	*kupu’	‘butterfly’	
Thai-C	p ^h i: sufa	‘butterfly’	
[Mon-Khmer data]			
BM, TM	kawa?	‘bat’	(ANLW?)
Pr	mə:k	‘bat’	
Pr	təkliw	‘butterfly’	(cf. Nic, Mal ‘bat’)
Nic	alo:âa	‘bat (<i>Miniopterus pusillus</i>)’	
Nic	kalu-mâwa	‘butterfly’	
Nic	mokngeaka	‘flying fox (<i>Pteropus Nic.</i>)’	
Asl-SkBl:	Sem. Jarum tawag	‘butterfly’; Söm ; Pang. Sam. tawag; Sem. K. Ken. <u>tauáng</u> ; Bes. K. L. <u>tawong</u> ; Sem. K. Ken. tabo ^g ng	‘dragon-fly’ (Skeat and Blagden 1966:551.B481)
Note also:			
Mkn-Lws	lolak	‘flying fox, bat’	
Mkn-KS	lolak	‘bat’	(cf. Nic ‘bat’)
Mkl-KY	?alə:?	‘Red Squirrel’	(poss. TLW)
UAN	*kalu’añ	‘bat’	
Thai-C	krarə:k	‘squirrel’	(ANLW?)
Thai-S	lə:k	‘Red Squirrel’	

Mkl *hãl/hêl* ‘at, to’ shows phonetic similarity with the Aslian Besis dialects, which Benjamin (1976:128) calls Mah Meri (MM). Note, for example, Asl-SkBl Bes. K. Lang. *ha* ‘at, in’, Bes. Malac. *ha* ‘to, with’, Bes. Sep. A. I. *ha* ‘to, towards’, Bes. Sep. A. I. *ha, hang* ‘with, against’ (Skeat and Blagden 1966:519.A178). In addition, the unexpected nasalized vowel of Mkl *hãl/hêl* ‘not’ could be plausibly explained as a trace from *-ŋ > -k > -?*.¹⁹ Compare, especially, Bes. Sep. A. I. *ha, hang* ‘with, against’ with Mkl *hãl/hêl*.

¹⁹ For a case that is similar, compare Ach ‘flower’ with Mkn-Mkl and Aslian data.

One of the most interesting putative MK loans is **PMM** ^ʔ*ʔəkəŋ* ‘mouth (human, animal)’ due to the possible correspondence between Mkn-Mkl *k* and Aslian *h* [cf. **Proto-Monic** **pa.ŋ*, **PNyK** **p_a.ŋ* ‘mouth (human, animal); opening; entrance; doorway’ **PMon** **pa_iŋ* ‘mouth, mouthful, quantifier for speech’ > BM, TM *paŋ*, NyK *pa.ŋ*; Asl-Benj: KS, KB, BD *haŋ*, CW *həŋ*, Asl-SkBl: SemBuk Max *hǎŋ* ‘mouth’, Sem Pa Max *hǎŋ* ‘mouth’.

On the possible affinity between **PMM** **ʔəkəŋ* ‘mouth’ and Nic *â-fàŋ/oal-fâŋ* ‘id.’, the likelihood of a connection is increased by the fact that the Southern Thai and Moklen people often employ [k^{hw}] where Central Thais use [f]. This possible correspondence is supported by the possible correspondence between Mkl-Lmp *kwa.p* ‘to yawn’ and Nic *fâp* ‘to yawn’.

Comparing **PMM** **ʔəkəŋ* and [Asl-SkBl] Sem Buk Max *hǎŋ* ‘mouth’, Sem Pa Max *hǎŋ* ‘mouth’ (Skeat and Blagden 1966:664.199) provides another feasible connection.

In regard to the possibility of secondary loan routes, some loans may have been borrowed indirectly by one or more of the mainland Austronesian languages. One possibility is primary MK influence on **PAC** or one of its daughter languages with secondary borrowing into **PMM**, Mkl, and/or Mkn from **PAC**, **PChm**, **PAch**, or their daughter languages. Other MK loans may have been borrowed indirectly into Mkn-Mkl via Thai.

If trans-peninsular trade was as important as most histories of the Thai-Malayan Peninsula claim, then we should be able to find linguistic correlations to the trade routes illustrated in Map 8. The lexical comparison that follows supports the probability of AN–MK trade along trans-peninsular river routes:

PMACM	<i>*buŋa[:]ʔ</i> (poss. <i>*buŋi[:]ŋ</i>)	‘flower’
PMM	<i>*buŋa:ʔ</i>	‘flower; country (kingdom)’
Mkn-Dng	<i>buŋa:ʔ</i>	
Mkn-Rn	<i>p^huŋa:ʔ</i>	‘country’
Mkn-KS	<i>p^huŋa:ʔ</i> [<i>p^hu.ŋa:ʔ</i>] [<i>b~p^h</i>]	
Mkn-Lws	<i>bunga</i>	
Mkn-Rw	<i>buŋaʔ</i>	
Mkl-KY	<i>muŋa:ʔ</i> [<i>mu.²¹.’ŋa:ʔ</i>] { <i>mu:ŋa:k⁸</i> }	‘flower; country’
Mkl-Lmp	<i>buŋa:ʔ</i>	
PAch	<i>*buŋãŋ</i> > Ach-Dur <i>buŋɤŋ</i> , Ach-Saw <i>buŋɔŋ</i> (-ŋ unexplained)	
PChm	<i>*buŋã</i> revised as <i>*buŋã:</i> (poss. <i>*buŋi:</i>)	‘flower’ > Rad <i>mŋa</i> , Jor <i>bəŋa</i> , Rog <i>buŋã</i> , Chm <i>pəŋu</i> [Rad-DB <i>mŋa:</i> ‘fleur’, Chm-Mou <i>pŋu:</i> ‘fleur’];
PMal	<i>*buŋa(?)</i> > Mal, UL <i>buŋa</i>	
PMP	<i>*buŋa</i> (poss. <i>*buŋaʔ</i>)	
POL-MUL	<i>*bu:ŋa:ʔ</i>	

- PPh** *bú:dak
PPh bú:ŋah ‘fruit’
PNPh *sa:buŋ
 [Ray (Borneo)] ‘country’ **I**: Lemanak benoa/menoa, Singhi binoa; **II**:
 Bolongan b’nua; **III**: Bekiau bagun, Bisaya pagun;
UAN *buŋa‘
UAN *majaŋ ‘flower (young spike of)’
 [Mon-Khmer] **Proto-Monic** *pkaaw, **PNyK** *kaaw ‘flower’, **PMon** *kaŋ
 ‘flower’
 Asl-Benj: Ks buŋa?, KB buŋa?, Je boŋe?, BD boŋa?, Mt boŋa?, Tm boŋa:?,
 SmI boŋa:?, SmII boŋa?, MM, boŋa? ‘flower’ [prob. ANLW, poss.
 from **PMACM** or one of its branches]

Trans-peninsular trade contact may explain the extensive Austronesian influence in Asl-N, Asl-C, and Asl-S words for ‘flower’. Routes 7 through 10 of Map 8 (Larish 1999, Book 1, p. 439) connect this widespread loan. Since it is widely distributed, this indicates that it may be an early loan. The borrowing in Aslian languages could, therefore, have its source in **PMACM** or one of its daughters. Geographically, the southern or central Thai-Malayan Peninsula is the most logical place for the **PMACM** homeland, for, like spokes in a wheel, here we find the only geographic hub connecting the present-day locations of **PMACM**’s daughter languages. The Aslian ‘**flower**’ data suggests that the first breakup of **PMACM** occurred when the various branches of the subgroup moved into different estuaries (e.g., routes 7 and 8 on Map 8, Larish 1999, Book 1, p. 439).

Returning to trade items, let us examine another commodity that was probably commonly exchanged in AN-MK interaction:

- PMACM** *sira:(?) ‘salt, n.’
PMM *c^hεla:ʔ/*sεla:ʔ ‘salt, n.’ (< **Pre-MM** *sira:?)
 Mkn-KS c^hεla? ‘salt, n.’
 Mkn-Lws chéla ‘salt’ (more recent data suggests è)
 Mkn-Rw sela? ‘salt, n.’
 Mkl-KY c^hεla:ʔ ‘salt’ [c^hε.²¹.!la:ʔ⁴⁵] {chε:la:k}
 Mkl-Lmp c^hεla:ʔ ‘salt’
PACM *sira[:] ‘salt, n.’ > **PAC** *sira[:]
PChm *sara revised as *sara: ‘salt’ > Rad, Jor hra, Rog saia, Chm sara
 [Rad-DB hra: ‘sel’, Chm-Mou sara: ‘sel’]

- PACM** *masin ‘salty, salted; briny, brackish’²⁰ > **PAC** *masin
(cf. SubS, SubSc masin ‘salt’ in PhML)
- Ach-Shr masen ‘salt, briny’
- PChm** *masin ‘salted’ > Rad msin, Jor masin/masin, Rog
masit, Chm miŋtin [Chm-Mou muthin ‘brine’, Jar məsin, Rög mösin]
- PMal** *garəm > Mal garam, UL garap [cf. Brunei Malay sira]
- PMP** *qasiya poss.?(qa-)siyá[:](?); **PMP** *timus ‘salt’
- PMP** *ma-qasin
- PPh** *qāsín ‘salt’ > Png, Tag asín
- [PhML ‘salt’] BilK, BilS, Tbl kahi?, numerous instances of ?asin and
?ahin, Mbl, MbKC timus, SubS, SubSc masin
- [Ray (Borneo)] **I**: Matu siah, Dali wai, Berawan usän, Kanowit siah, Tanjong
chia / Sibuyau garum, Lemanak garam / Quop garo, Sau galu; **II**:
Manyan rangi / Bologan garam, Tarakan asin, Sesayap masin, Sibuku
asin / Uma Baloi hia / Lepu Tau usən, Long Bangan osan; **III**: Kelabit
tuchu / Balait tucho / Padas usi / Kajaman siyã, Rejang usen / Sibop
uchən / Ukit jio, Bakatan ijo / Kadayan usän / Tagal masin / Dalit
usun, Tampasuk asin, Ida’an silan, Lanun timus, Buludupi tagai,
Buludupi mawsin, Brunei Malay sira
- [Purwo (Indo-C) ‘salt’]A: Ind garam, Snd uyah, Jav uyah, Mad buja, Bal uyah
[Purwo (Indo-C) ‘salt’]B: Ind garam, Snd—, Jav sarem, Mad —, Bal tasik
- Rej siləy
- Sulu a:sin
- UAN** *‘at’in ‘salt’
- Thai-C klua
- [Mon-Khmer] **Proto-Monic** *bʔur ‘salt’ > **PNyK** *phəʔur, **PMon** *bə
- BM, TM dəŋ, NyK pəʔur ‘salt’
- Pr lu(:)k/lo(:)k
- Khm ʔəmbɔl ‘salt, n.’
- Asl-SkBl Sem. Craw. Hist siyak, Sem. Plus siäk/siā‘, Sem. Ked. And. siyah,
Sem. Kedah siak

Note that the Aslian languages with probable An loans for ‘salt’ could not have borrowed these words from Malayic, since we find **PMal** *garəm ‘salt’ > Mal *garam*, UL *garap*. Note also that **PAC** *sira[:] ‘salt’ has no glottal final. Thus, these Aslian forms were probably borrowed from either **PMACM** *sira: (?) ‘salt, n.’ or **PMM** *c^hela:ʔ/*sela:ʔ ‘salt, n.’ (< **Pre-MM** *sira:ʔ).

²⁰ Blust (1994:40) proposes *masin ‘salty’ as a **PACM** morphological innovation.

Additional Aslian words with final *-k* include ‘wash clothes’ and ‘wet’ (Skeat and Blagden 755.W75). These words are significant, for Aslian final *-k* may be attributed to either **Pre-MM** or **PMACM** influence; they cannot be loans from **PACM** or one of its daughter languages, where $*q > h$ is expected.

3. Conclusion

The fact that AN-MK data cluster into specific semantic domains suggests a possible early trading relationship between the mainland-AN languages and one or more MK languages. The nature of the borrowing is complex, and in many cases, especially in core vocabulary, the relationship between Austronesian and MK is probably not diffusional at all, but one of remote common origin. The mainland-AN languages may retain more Austric word-stock than insular-AN because extended contact may have encouraged the retention of common vocabulary. Even in the cases where borrowing is suspected in Moken and Moklen, no consistent, discernible MK source can normally be pinpointed, suggesting early borrowing. Moken and Moklen share affinities variously with Aslian, Nicobarese, Monic, and Nyah Kur languages. In cases where no direct source can be determined, this may suggest that some words that appear to be loans are, in fact, the result of a genetic connection. Other possibilities are that the Mon-Khmer influence came through conquest, intermarriage, or trade—each implying a certain degree of bilingualism.

When Moken-Moklen and MK data are compared, especially Aslian data, lexical similarities suggest that contact existed between speakers of **PMM** and early peninsular MK speakers. Austronesian and Mon-Khmer-speaking groups largely exploited different domains, yet they were drawn to interact through barter or trade. If in fact the **PMM** or **Pre-MM** participated in trans-peninsular trade, they would have frequently interacted with interior peoples.

Appendix 1.1: Comparanda between PAn (Blust 1997), PST (Coblin 1986), and PTB (Matisoff [=JAM] 2003)

1	Proto-Forms	Glosses	Sources
PAn	*ka-wiRi	left side	2/200 Blust (1997)
PTB	*bi(y) *b ^w ay	left	JAM 03:584R, 585R, 657L
2	Proto-Forms	Glosses	Sources
PAn	*ka-wanaN	right side	3/200 Blust (1997)
PTB	*g- ra~ *g- ya	right hand right side	JAM 03:609C, 664R
3	Proto-Forms	Glosses	Sources
PAn	*qaqay	leg/foot	4/200 Blust (1997)
PTB	*kaŋ~ *keŋ	leg/foot/stem/stalk	JAM 03:595C, 657L
4	Proto-Forms	Glosses	Sources
PAn	*sakay	walk/go	5/200 Blust (1997)
PST	*sywjaj	go	Coblin (1986:86)
OC	*gwjag	go	Coblin (1986:86)
PTB	*s-ka-y	go/stride	JAM 03:594C, 652R
^a . PAn *sakay and PTB *s-ka-y present striking comparanda.			
5	Proto-Forms	Glosses	Sources
PAn	*zalan	road	6/200 Blust (1997)
PTB	*lam	road	JAM 03: 559L, 665L
6	Proto-Forms	Glosses	Sources
PAn	*CuqelaN	bone	15/200 Blust (1997)
PTB	*g-rus ^a *s-rus *m-rus	bone	JAM 03: 102/611C/650R
^a . Cf. PLB *rəw ² 'bone'			
7 & 8	Proto-Forms	Glosses	Sources
PAn (7)	*-ajem	think	21A/200 Blust (1997)
PST	*niəm	think	Coblin (1986:148)
OC	*niəmh	think	Coblin (1986:148)
PTB	*s-nyam	think	JAM 03:605C, 671R
PAn (8)	*nemnem	think	21B/200 Blust (1997)
PST	*nyəm	think	Coblin (1986:148)
OC	*njəmx	think	Coblin (1986:148)
PTB	*s-nyam	think	JAM 03:605C, 671R

9	Proto-Forms	Glosses	Sources
PAn	*daRaq	blood	23/200 Blust (1997)
PST	*khrjak	red/blood	Coblin (1986:123)
OC	*khrjak	red/blood	Coblin (1986:123)
PTB	*ts yak	red/blood/gold	JAM 03:617R, 641R
	^{a.} cf. dz(y)ak~*ts(y)ak ‘drip/drop’ (JAM 03:617R)		
10	Proto-Forms	Glosses	Sources
PAn	*liqeR	neck	25/200 Blust (1997)
PTB	*s-ke-k	neck/neck shaped	JAM 03:595C, 659R
11	Proto-Forms	Glosses	Sources
PAn	*nipen ^a	tooth	31/200 Blust (1997)
PTB	*ŋa	tooth	JAM 03:605R, 672C
	^{a.} Cf. PAT *(N)Gi(m)pan ‘tooth’ (PKB 1990:255, 1975:411)		
12	Proto-Forms	Glosses	Sources
PAn	*tanek	cook	39/200 Blust (1997)
PTB	*k lak	cook/boil	JAM 03:595R, 645C
PAn	*kaRat ^a	bite	41/200 Blust (1997)
PTB	*ts at	bite down on	JAM 03:616R, 641C
	^{a.} PAn k<aR>at?		
13 & 14	Proto-Forms	Glosses	Sources
PAn (13)	*susu	breast	18/200 Blust (1997)
PTB	*tsyup~	suck/kiss/breast/	JAM 03:618L/642C
	*tsyip	milk	
PAn (14)	*sepsep	suck	42/200 Blust (1997)
PST	*tsop~	suck	Coblin (1986:144)
	*dzop		
OC	*tsəp	suck	Coblin (1986:144)
PTB	*tsyup~	suck/kiss/breast/	JAM 03:618L/642C
	*tsyip	milk	
PTB	*dzyut	suck/kiss/breast/	JAM 03:589C/642C
		milk	
	^{a.} PAn susu ‘breast’ and PAn sepsep ‘suck’ are probably members of the same word family (allofams, see Matisoff = JAM 03). Although the TB semantics have a wider range, the phonosemantic overlap between two etyma in PAn (breast and suck) and their affinities to PTB support the PASN-Hyp . The PAK data in Appendix 3 present another case for possible cognicity. PTB has a wealth of variant forms (allofams) that are not presented here (Cf. JAM 03:642C).		
15	Proto-Forms	Glosses	Sources
PAn	*maCa	eye	45/200 Blust (1997)
PST	*myikw	eye	Coblin (1986:76)
OC	*myəkw	eye	Coblin (1986:76)
PTB	*s-myak~	eye	JAM 03:602L, 649C
	*s-myik		
16	Proto-Forms	Glosses	Sources

PAn	*Rumaq	house	61/200 Blust (1997)
PTB	*k-yum *k-yim	house	JAM 03:620C, 654R

17	Proto-Forms	Glosses	Sources
PAn	*zaRum ^a	needle	68/200 Blust (1997)
PTB	*k-ram~ *k-rap	needle	JAM 03:609R, 659R

^a When I find Austro-Asian vowel correspondence between *i*, and/or *u*, and/or *a*, I hypothesize that they may reflect **PAsn** **i* (see Larish 1999, p. 301 for related discussion).

18	Proto-Forms	Glosses	Sources
PAn	*panaq	shoot	70/200 Blust (1997)
PTB	*m-p uk	shoot	JAM 03:607R, 667L

^a The unfixed **PTB** form supports Hayes's (2000) infixed analysis: *p<an>aq (see Reid 2005, Table 2).

19	Proto-Forms	Glosses	Sources
PAn	*Cakaw	steal	73/200 Blust (1997)
PST	*rkhuγ	rob	Coblin (1986:126)
OC	*khugh	rob	Coblin (1986:126)
PTB	*r-kəw ^a *r-kun *r-kut	steal/thief	JAM 03:595R, 669C

^a Coblin (1986:126) reconstructs PTB **r-kuw*.

20	Proto-Forms	Glosses	Sources
PAn	*taRaq	cut (wood)	78/200 Blust (1997)
PTB	*ts yat~ *m rak~ *b rat~ *p rat	break/cut cut/tear cut apart/cut open	JAM 03:617R, 646L JAM 03:602C, 646L JAM 03:585L, 646L

21	Proto-Forms	Glosses	Sources
PAn	*Cazem?	sharp	81/200 Blust (1997)
PST	*slyam	sharp	Coblin (1986:131)
OC	*sjam	sharp	Coblin (1986:131)
PTB	*s-ryam	sharp	JAM 03:612L, 666R

22	Proto-Forms	Glosses	Sources
PAn	*Ribawa	swell	87/200 Blust (1997)
PTB	*b wap	swell up/swollen/ stout/calf of leg	JAM 03:585R, 671L

23	Proto-Forms	Glosses	Sources
PAn	*baliw beli	buy	91/200 Blust (97)
PTB	*m-lay	change/exchange/	JAM 03:599C, 643C

	*r-ley	buy/barter	
24	Proto-Forms	Glosses	Sources
PAn	*tutuh	to pound	93/200 Blust (1997)
PST	*khuk ^a	strike/knock	Coblin (1986:142)
OC	*khugx	knock	Coblin (1986:142)
	*khugh		
PTB	*r- tuk	strike	JAM 03:615R, 670L
	^a . I have generally avoided onomatopoeic words, but I have included this set to present instances of <i>t</i> to <i>k</i> and <i>-h</i> to <i>-k</i> .		
25	Proto-Forms	Glosses	Sources
PAn	*qiCeluR	egg	98/200 Blust (1997)
PTB	*t wəy	water/egg/spit	JAM 03:616L, 648R
	*ʔu	egg/sit on eggs	JAM 03:583C, 648R
26	Proto-Forms	Glosses	Sources
PAn	*SimaR ^a	fat/oil	104/200 Blust (97)
PST	*s ayw	fat/grease	Coblin (1986:77)
OC	*s agw	fat/grease	Coblin (1986:77)
PTB	*tsi l	fat/grease/oil	JAM 03:612R, 617L,
	*s a:w		650L
PTB	*ts ow-s	fat/omentum	JAM 03:617 650L
	^a . Reid (2005, Table 2) presents Hayes's (2000) reanalysis as *S<im>aR.		
27	Proto-Forms	Glosses	Sources
PAn	*Su laR	snake	106/200 Blust (97)
PTB	*s-b/m-ru:l ^a	snake	JAM 03:611C, 668L
	^a . Cf. PLB *m-r-wəy ¹ 'snake'		
28	Proto-Forms	Glosses	Sources
PAn	*kakaCu	spider	110/200 Blust (97)
PTB	*kaŋ~	spider/spin	JAM 03:595L, 668R
	*waŋ		
29	Proto-Forms	Glosses	Sources
PAn	*Sikan	fish	111/200 Blust (97)
PST	*ŋyay	fish	Coblin (1986:80)
OC	*ŋyag	fish	Coblin (1986:80)
PTB	*s-ŋya	fish	JAM 03:606C, 650R
30 & 31	Proto-Forms	Glosses	Sources
PAn (30)	*qasiRa	salt	125A/200 Blust (97)
PTB	*la	salt	JAM 03:599L, 665R
	*m-t(s)i	salt/yeast	JAM 03:617L, 665R
	*tsa	salt	JAM 03:616C, 665R
PAn (31)	*timus ^a	salt	125B/200 Blust (97)
PTB	*m-t(s)i	salt/yeast	JAM 03:617L, 665R

*tσα salt JAM 03:616C, 665R
^{a.} **Pre-An** *ts can possibly be reconstructed here based on internal evidence: *si- in 125A & *ti- in 125B <? *tsi).

32 & 33	Proto-Forms	Glosses	Sources
PAn (32)	*bulaN	moon	129A/200 Blust (97)
PTB	*s/g-la	moon/month	JAM 03:599L, 659L
PAn (33)	*qiNaS	moon	129B/200 Blust (97)
PTB	*s-ŋ ^w (y)a-t	star/moon	JAM 03:606C, 659L

34	Proto-Forms	Glosses	Sources
PAn	*quzaN	rain	133/200 Blust (97)
PST	*rywjay	rain	Coblin (1986:122)
OC	*gwjagx	rain	Coblin (1986:122)
PTB	*g-wa~ *s-wa *r-wa	rain	JAM 03:618C, 663R

35	Proto-Forms	Glosses	Sources
PAn	*deRuŋ(?)	thunder	134/200 Blust (97)
PTB	*m-b r uŋ ^a ~ *m-b r uk	thunder/dragon	JAM 03:585C, 672L

^{a.} This may be disassimilation: d → b/_r?

36	Proto-Forms	Glosses	Sources
PAn	*likaC	lightning	135/200 Blust (97)
PTB	*gle:k ^a	thunderbolt/ lightning	JAM 03:590R, 657C

^{a.} possible metathesis

37	Proto-Forms	Glosses	Sources
PAn	*bali	wind	136/200 Blust (97)
PTB	*g-ləy ^a	wind (n.)	JAM 03:600L, 675L

^{a.} əy < *i: ?? For similar changes in mainland Austronesian languages, see Larish (1999:395-402).

38	Proto-Forms	Glosses	Sources
PAn	*CuNuh	burn	144/200 Blust (97)
PTB	*duk ^a	burn/kindle	JAM 03:587C, 643L

^{a.} Note -h may correspond to -k.

39	Proto-Forms	Glosses	Sources
PAn	*qebel	smoke	145/200 Blust (97)
PTB	*kəw-n/t	smoke	JAM 03:595C, 668L

40	Proto-Forms	Glosses	Sources
PAn	*CeŋeN	black	147/200 Blust (97)
PTB	*tyaŋ	black/dark	JAM 03:616C, 641R

41	Proto-Forms	Glosses	Sources
PAn	*ma-puNi	white	148/200 Blust (97)
PTB	*p lu	white/silver/money	JAM 03:607L, 674R

	Proto-Forms	Glosses	Sources
42			
PAn	*ma-taNah(?)	red	149/200 Blust (97)
PST	*khrjak	red/blood	Coblin (1986:123)
OC	*khrjak	red/blood	Coblin (1986:123)
PTB	*tsyak	red/blood/gold	JAM 03:617R, 664L
	*dzya-n	blush/red	JAM 03:588R, 664L
	*t(y)a-n	red	JAM 03:616C, 664L
	*n(y)a-n	blush/red	JAM 03:605C, 664L

^a. These are striking comparanda. The **PST** and **PTB** reflexes of a possible **PAn** form suggest that **Pre-An** *N may reflect a consonant cluster. In addition, here we find additional possible correspondences between **PAn** *t-* and **PST** *k-* and **PAn** *-h* and **PST** *-k*.

	Proto-Forms	Glosses	Sources
43			
PAn	*inaduq(?)	long (objects)	155/200 Blust (97)
PTB	*duŋ~	long/length	JAM 03:587C, 657R
	*tu:ŋ		

	Proto-Forms	Glosses	Sources
44			
PAn	*ma-NiSepis	thin (materials)	156/200 Blust (97)
PTB	*ly ap~	thin/flat/flat object	JAM 03:599R, 671R
	*l ep~		

	Proto-Forms	Glosses	Sources
45			
PAn	*ma-kaSepal	thick	157/200 Blust (97)
PTB	*r-tas	thick/solid/coarse	JAM 03:614R, 671R

	Proto-Forms	Glosses	Sources
46			
PAn	*ma-tuqaS	old (people)	162/200 Blust (97)
PTB	*b-gres	old	JAM 03:591L, 660R
PTB	*r-ga	old	JAM 03:589R, 660R

	Proto-Forms	Glosses	Sources
47			
PAn	*qalejaw	day	168/200 Blust (97)
PST	*nyi?	sun/day	Coblin (1986:145)
OC	*njit	sun/day	Coblin (1986:145)
PTB	*nəy ^a	sun/day	JAM 03:604R, 646C

^a Coblin (1986:145) reconstructs PTB *nyiy.

	Proto-Forms	Glosses	Sources
48			
PAn	*i-taqaS	above	175/200 Blust (97)
PTB	*l-tak	ascend/lift/raise/top	JAM 03:614C, 640L

	Proto-Forms	Glosses	Sources
49			
PAn	*ma-azaNih	near	179/200 Blust (97)
PST	*nyiy	near	Coblin (1986:111)
OC	*njirx	near	Coblin (1986:111)
PTB	*s-ney	near	JAM 03:604C, 659R
	*s-na:y		

50	Proto-Forms	Glosses	Sources
PAn	*ma-dawIN	far	180/200 Blust (97)
PTB	*g-wəy-n	far	JAM 03:619C, 649R
51 & 52	Proto-Forms	Glosses	Sources
PAn (51)	*esa	one	197A/200 Blust (97)
PTB	*t(y)ak ^a	one/only	JAM 03:616C, 660R
PAn (52)	*isa	one	197B/200 Blust (97)
PTB	*ʔit ^b	one	JAM 03:583C, 660R

^a. Here is another case of phonosemantic overlap between two sets of etyma. For another, see breast/suck.

^b. Cf. **PAK** *ʔitsa ‘one’ (PKB 1990:224-5)

Appendix 1.2: PMP (Blust 1997) vs. PTB (Matisoff 03) Comparanda

53	Proto-Forms	Glosses	Sources
PMP	*qabuk/qapuk	dust	11/200 Blust (1997)
PTB	*mu:k	detritus/dust	JAM 03:602R, 648L
54	Proto-Forms	Glosses	Sources
PMP	*qinep	lie down	49/200 Blust (1997)
PTB	*s-yip	sleep/put to sleep/	JAM 03:620C, 667R
	*s-yup	conceal/hide (v.)	
55	Proto-Forms	Glosses	Sources
PMP	*tudan ^a	sit	51/200 Blust (1997)
PTB	*m-duŋ/k~	sit	JAM 03:587C, 667C
	*m-tuŋ/k		

^a. Note **UAN** *ɖukɖuk ‘to sit’ > Moklen dɔ:k ‘to sit’

56	Proto-Forms	Glosses	Sources
PMP	*si laq	split	80/200 Blust (1997)
PTB	*(t)si: t~	split	JAM 03:588C, 669L
	*(d)zi: t~	split	
	*dzi k	split/mince	
57	Proto-Forms	Glosses	Sources
PMP	*buka	to open	92/200 Blust (1997)
PTB	*s-bu	bud/open	JAM 03:585C, 660R
PTB	*m-ka	open/opening/ mouth/door	JAM 03:594C, 660R
58	Proto-Forms	Glosses	Sources
PMP	*bulu	feather	99/200 Blust (1997)

PTB	*g-mul	hair (body)	JAM 03:602R, 650L
59	Proto-Forms	Glosses	Sources
PMP	*kabut	fog/mist	132/200 Blust (97)
PTB	*r/s-mu:k~	overcast/foggy/dark/ sullen	JAM 03:603L, 646C
	*mu:ŋ		
60	Proto-Forms	Glosses	Sources
PMP	*ma-kunij	yellow	150/200 Blust (97)
PTB	*hwaŋ	shine/bright/yellow	JAM 03:593R, 675R
PTB	*hwa:r~	white/yellow/bright/ shine	JAM 03:593R, 675R
	*yar		
PTB	*hwa:r	fire/burn/shine/white	JAM 03:593R, 675R
61	Proto-Forms	Glosses	Sources
PMP	*kepit	narrow	158/200 Blust (97)
PTB	*gyap	narrow/crowded	JAM 03:592C, 659R
62	Proto-Forms	Glosses	Sources
PMP	*ma-sakit	sick, painful	160/200 Blust (97)
PTB	*tsa-t	hot/hurt/pain/ill	JAM 03:616C, 661C

Appendix 2: Expanded Comparanda

Proto-Lng	Proto-Form	Gloss	Source
PAT-75	*k[i]ᵛ i/(k[i]ᵛ i)	tickle/armpit	PKB 1975:410
PTB	*g- li	armpit/tickle	JAM 03:600L, STC#265 639R
PAT-75	*[(m)po(ŋ)]kor	behind/back buttocks	PKB 1975:230
PTB	*r-tul~ *r-til	dull/buttock/heel rounded part	2003:615R, 643L
PAn	*likud	back	13/200 Blust (1997)
PAT-75	*(qa/ka)ᵛ i[kuz]	back/behind	PKB 1975:223-4
PTB	*s-nuk~ *s-nuŋ	back/behind/after	2003:605L
PTB	*m-kal~ *s-ga:l	kidney/small of back/loins	2003: 12, 590L, STC#12 640C
PAT-75	*[ku]ᵛ uᵛ	back ^a	PKB 1975:223
PTB	*m-g lun	kidney	JAM 03:73
OC	d̪iən	small of the back, reins	JAM 03:73

^a Although the semantics appear dissimilar (i.e., **PAT** ‘back’ and **PTB** ‘kidney’), the TB data in Benedict (1972:18, STC#12) demonstrate overlap between kidneys/loins/small of the back/back. Moreover, the possible phonemic correspondence between **PAT** and **PTB** is striking.

Proto-Lng	Proto-Form	Gloss	Source
PAT-75	*(q/)u(m)pak	bark, rind, skin, pod, husk	PKB 1975:225
PAJ	*kaba	skin	PKB 1990:242
PTB	*k wa(:)k	skin	JAM 03:667C
PTB	*kok~	outer covering/bark (n.)/rind/skin	JAM 03:596L/640C
PTB	*r-kwa(:)k	outer covering	JAM 03:596L/640C
Proto-Lng	Proto-Form	Gloss	Source
PAT-75	*(m)baᵛ i	exchange, change, buy, sell	PKB 1975:282
PTB	*b-rey	buy/barter	JAM 03:610L/640C
PTB	*r-ley~ *g/m/s-lay	change/exchange/buy barter	JAM 03:599C/640C

Proto-Lng	Proto-Form	Gloss	Source
PAT-75	*()(m)praŋ	bee	PKB 1975:229
PTB	*s-braŋ	fly (n.)/bee	JAM 03:585L/641L
Proto-Lng	Proto-Form	Gloss	Source
PAT-75	*[kə]mpuŋ ^a	belly	PKB 1975:230
PTB	*pu:k~	belly/cave	JAM 03:607R/641L
	*bu:k		
PMonic	*bu[u] ŋ	belly (of humans, animals, jars...)	Diffloth 1984:98
PAT-75	*[i](ŋ)kuŋ (m)b/[i](ŋ)kuŋ	bend/bent, arched, crooked	PKB 1975:231
	*[i] ŋkuk		
	*(m)b/[i] ŋkuk		
	*b/u ŋkuk		
	*uguk		
PTB	*gu(:)k ~*m-ku(:)k	crooked/bent/knee/ angle	JAM 03:591R/641L
PTB	*ʔuk~ *kuk	crooked/bent/knee/ angle/return/back	JAM 03:583C/641L
PAT-75	*(q/)ʔay[a]	big, long	PKB 1975:233
PLB	*k-ri(y) ²	big	JAM 03:611L
PAT-75	*(ts)[i]rat	bind, tie, knot, squeeze	PKB 1975:233
PTB	*k(y)it/k~ *g(y)it/k	tie/bind	JAM 03:592R/641C
PTB	*g rak	cord/bind/tie	JAM 03:591L/641C
PAT-75	*ntsa[a]mu[ʔ] ^a	blood	PKB 1975:235
PTB	*tšyak	red/blood/gold	JAM 03:617R/641R
PTB	*ts(y)ak~ *dz(y)ak	drip/drop (n.)	JAM 03:617R
PAn	*daRaq	blood	23/200 Blust (1997)
^a . Cf. P-Monic	*chim ‘blood’ (Diffloth 1984:103)		
PAT-75	*[ts]ə(m)put (/(m)put)	blow, wind	PKB 1975:236
PTB	*s-mut	blow	JAM 03:641R

PAT-75	*[da]Giŋ	body/flesh	PKB 1975:238
PTB	*guŋ	body	JAM 03:592L/642L
PTB	*s-kəw	body/corpse	JAM 03:595C/642L
PAT-75	*ts[a]ŋa	fork, branch, groin	PKB 1975:297
PAT-75	*[da]Nqa	branch	PKB 1975:240-1
	*[(n)da](N)qa/n		
	*[da](ŋ)ka		
	*[d]aka/n		
PTB	*s-ka:k	fork/branch	JAM 03:594R
PTB	*ku:ŋ	tree/branch/stem	JAM 03:597R
PAT-75	*nu[h]/(nu[h])	breast	PKB 1975:242
PTB	*nəw	breast/milk	JAM 03:642C
PAK-90	*tsitsi	breast	PKB 1990:173
PTB	*tsyip~	suck/kiss/breast/	JAM
	*tsyup ^a	milk	03:618L/642C
PAn	*susu	breast	18/200 Blust (1997)
^a . For additional variant forms, see JAM (2003:642C). Many PTB etyma appear related: PTB *g-ts(y)i-t/n~*zəy ‘urine’, *m-ts(y)il~*til ‘spit/spittle/saliva’, suggesting that Pre-TB *tsyi- may have meant ‘liquid’ (JAM 03:617-8).			
PAT-75	*[ta]pats(/pats)	sweep, dust, shake, broom	PKB 1990:173
PTB	*py(w)ak	sweep/broom	JAM 03:609L/642R
PAT-75	*g[a]lak ^a	burn/blaze	PKB 1975:244
PTB	*duk	burn/kindle	JAM 03:587C/643L
PTB	*g- duk ^b	daytime/noon	JAM 03:587C/643L
PLB	*?-duk ^L	burn/kindle	JAM 03:587C/643L
PLB	*?-gaŋ ^l	roast/toast/burn/be dry	JAM 03:590L
^a . Could the PAT form have an <-al-> infix?			
^b . This set illustrates the methodological problems inherent to this research. Given a possible g-d-l correspondence (Cf. Larish 1999:173, Table 3.2), the comparanda can vertically aligned in a number of different ways.			
PAT-75	*(q/)(n)tu aŋ	bone ^a	PKB 1975:238
PAJ-90	*[ts,tš]aRap	offal	PKB 1990:224
PTB	*g-ra	fishbone/spine	JAM 03: 609L/650R

OC	*gli _o	spine	JAM 03:173
PTB	*s-rus *m-rus *g-rus ^b	bone	JAM 03: 102/611C/650R
OC	*kwæt	bone	JAM 03:465
PAn	*CuqelaN		Blust (1997:43)
PMP	*tuqəlán *tuqəlán	bone (condylar)	
PPh	*tuqǎlán	bone	Larish (99:20)
PTai	*ʔdl/ruok D1L	bone	Li (1977:129, 267)
PMK	*jʔaaŋ	bone ^c	

^a. The data from **PMY** *tshuŋ (Benedict 1975:239) supports the reconstruction of initial *tʃ.

^b. cf. PLB *rəw² ‘bone’

^c. **PMK** *jʔaaŋ, **Proto-Monic** *j[ɫ]uut, **PNyK** *chəlɥut, **PMon** *cɥt

PAT-75	*[q]aplay	liver, entrails, heart (emotions), mind	PKB 1975:332
PTB	*b-ka-n	bitter/liver	JAM 03:594C, STC#8 657R
OC	kân	liver	JAM 03:451
PAn	*qaCay	liver	17/200 Blust (1997)

PAT-75	*(m)ba/ba	carry (on back)	PKB 1975:246	
PTB	*ba	carry (on back)	JAM 2003:643C	
PAJ-90	*ga(ŋ)ki ^a	crab	PKB 1990:178	
OJpn	*[-n-kani	crab	Martin 1987:437	
PTB	*d-k(y)an	crab	JAM 2003:645R	STC#51
PMonic	*kntaam	fresh-water crab	Diffloth 1984:77.N52	
^a Benedict cites PAn *gaki ‘crab’ based on PNPh data and Kankanay <i>gaki</i> .				
PAJ-90	*talak	cook/roast	PKB 1990:177	
PAn	*tanek	cook	39/200 Blust (1997)	
PTB	*k lak	cook/boil	JAM 03:595R, 645C	
Note the striking possible initial *t to *k correspondence between PAJ and PTB.				
PAT-75	*(n)dza[a](m)bot	hair, beard	PKB 1975:306	
PTB	*tsam~ *sam	hair (head)	JAM 2003:616R/653R	STC#73
PAT-75	*[gu]mul	hair (body)/eyebrow	PKB 1975:308	
PTB	*g-mul	hair (body)	JAM 2003:602R/653R	STC#2
PTB	*s-mul~ *s-mil*s-myal	hair (body)/fur feather	JAM 2003:603L/653R	

References not found in Larish (1999)

- Benedict, Paul K. 1975. *Austro-Thai Language and Culture with a Glossary of Roots*. Human Relations Area Files.
- Benedict, Paul K. 1976. *Austro-Thai and Austroasiatic*. In: Philip N. Jenner, Laurence C. Thompson, and Stanley Starosta, eds., *Austroasiatic Studies, Part I*, pp. 1-36. Honolulu: University of Hawai'i Press.
- Benedict, Paul K. 1990. *Japanese/Austro-Tai*. Ann Arbor: Karoma.
- Benjamin, Geoffrey. 1976. *An outline of Temiar grammar*. In: Philip N. Jenner, Laurence C. Thompson, and Stanley Starosta, eds., *Austroasiatic Studies, Part I*, pp. 129-188. Honolulu: University of Hawai'i Press.
- Blust, Robert A. 1997. *Subgrouping, circularity, and extinction: Some issues in Austronesian comparative linguistics*. *Selected Papers from the Eighth International Conference on Austronesian Linguistics*, Taipei: Academia Sinica.
- Diffloth, Gérard. 1984. *The Dvaravati Old Mon language and Nyah Kur*. *Monic Language Studies, Volume 1*. Bangkok: Chulalongkorn University Printing House.
- Diffloth, Gérard. 1994. *The lexical evidence for Austric, so far*. Paper presented at the Conference on Asia-Mainland/Austronesian Connections. Honolulu, May 10-13, 1993. *Oceanic Linguistics* 33:309-22.
- Hayes, La Vaughn H. 1992. *On the track of Austric: Part I. Introduction*. *Mon-Khmer Studies* 21.143-77.
- Hayes, La Vaughn H. 1997. *On the track of Austric: Part II. Consonant mutation in early Austroasiatic*. *Mon-Khmer Studies* 27.13-44.
- Hayes, La Vaughn H. 1999. *On the track of Austric: Part III. Basic vocabulary comparison*. *Mon-Khmer Studies*, 29.1-34.
- Jenner, Philip N., Laurence C. Thompson, and Stanley Starosta, eds. 1976. *Austroasiatic Studies, Part I and Part II*. *Oceanic Linguistics Special Publication No. 13*. Honolulu: University of Hawai'i Press. 1343 pp.
- Larish, Michael D. 1991. 'The special relationship between Moken, Acehnese, Chamic and Mon-Khmer: Areal influence or genetic affinity?'

- Unpublished paper presented at the *Sixth International Conference on Austronesian Linguistics*. Honolulu: University of Hawai'i.
- Larish, Michael D. 1993. 'Who are the Moken and Moklen on the Islands and Coasts of the Andaman Sea?' in *Pan-Asiatic Linguistics: Proceedings of the Third International Symposium on Language and Linguistics*, Chulalongkorn University, January 8-10, 1992. Volume III:1305-19. Bangkok: Chulalongkorn University Printing House.
- Larish, Michael D. 1997. 'Moklen-Moken phonology: Mainland or insular Southeast Asian typology?', in C. Odé and W. Stokhof (eds.) *Proceedings of the Seventh International Conference on Austronesian Linguistics*, 125-50, Amsterdam: Rodopi.
- Larish, Michael D. 1999. *The Position of Moken and Moklen in the Austronesian Language Family*, Ph.D. dissertation, University of Hawai'i at Mānoa. Ann Arbor, Michigan: UMI Dissertation Services.
- Larish, Michael D. 2005. Moken and Moklen. In: *The Austronesian Languages of Asia and Madagascar*, Chapter 18, pp. 513-33. Routledge Language Family Series. Alexander Adelaar and Nikolaus P. Himmelmann (eds.). London: Routledge.
- Li, Charles N. and Sandra A. Thompson. 1981. *Mandarin Chinese: A functional reference grammar*. Berkeley: UC Press.
- Martin, Samuel E. 1987. *The Japanese language through time*. New Haven: Yale University Press.
- Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction*. Berkeley: UC Press.
- O'Grady, William, John Archibald, Mark Aronoff, and Janie Rees-Miller. 2005. *Contemporary linguistics: An introduction*. 5th edition. Boston: Bedford/St. Martin's.
- Reid, Lawrence A. 2005. The current status of Austric: A review and evaluation of the lexical and morphosyntactic evidence. In: *The peopling of East Asia: Putting together archaeology, linguistics, and genetics*, ed. by Laurent Sagart, Roger Blench, and Alicia Sanchez-Mazas, 134-162. London: Routledge Curzon.
- Ruhlen, Merrit. 1991. *A guide to the world's languages*. Vol. 1: Classification. (2nd ed.) London: Edward Arnold.
- Sagart, Laurent, Roger Blench, and Alicia Sanchez-Mazas, eds. 2005. *The peopling of East Asia: Putting together archaeology, linguistics, and genetics*. London: Routledge Curzon.

- Starostin, Sergei. 1992. Methodology of long-range comparison. In: Nostratic, Dene-Caucasian, Austric and Amerind: Materials from the first international interdisciplinary symposium on language and prehistory, Ann Arbor, 8-12 November, 1988, ed. by Vitaly Shevoroshkin. Bochum: Universitätsverlag Dr. Norbert Brockmeyer.
- Thurgood, Graham. From Ancient Cham to modern dialects: Two thousand years of language contact and change. Oceanic Linguistics Special Publication No. 28. Honolulu: UH Press.
- Trask, R. L. 2000. The dictionary of historical and comparative linguistics. Chicago: Fitzroy Dearborn Publishers.

The preceding document was presented at the Tenth International Conference on Austronesian Linguistics (10-ICAL). To properly reference this work, please use the following format:

<LastName>, <FirstName>. 2006. <PaperTitle>. Paper presented at Tenth International Conference on Austronesian Linguistics. 17-20 January 2006. Puerto Princesa City, Palawan, Philippines.
<http://www.sil.org/asia/philippines/ical/papers.html>

For other papers that were presented at 10-ICAL, please visit <http://www.sil.org/asia/philippines/ical/papers.html>.